THEORY OF REGULATORY COMPLIANCE

DR RICHARD FIENE

A Treatise on the Theory of Regulatory Compliance

Richard J. Fiene[a,*]

[a]*Edna Bennett Pierce Prevention Research Center, Pennsylvania State University, 305 Templar Drive, Elizabethtown, Pennsylvania 17022*

Abstract

This treatise provides some insights into certain assumptions related to regulatory compliance and the implications for regulatory researchers and policy-makers for the future development of rules and regulations. Once regulatory compliance decision making moves from requiring full compliance with all rules to a substantial regulatory compliance decision making approach, the measurement and monitoring systems employed to assess programs and facilities change dramatically.

Keywords: regulatory compliance, risk assessment, key indicators, licensing, monitoring, measurement

1. Introduction

Regulatory compliance is a sub-discipline within regulatory science that focuses on measurement, monitoring systems, risk assessment, and decision making based on regulatory compliance scoring. Regulatory compliance is dominated by nominal scale measurement, that is, either a facility is in or out of compliance with specific rules. There is no middle ground with regulatory compliance as there is with most quality measurements, which are generally made on an ordinal scale. However, some regulators feel that certain regulations are not or should not be subjected to nominal measurement.

A factor with regulatory compliance data is that they generally follow a very skewed frequency distribution, which limits analyses to non-parametric statistics. Because of the skewed data distribution, dichotomization of data is warranted, given the lack of variance in the regulatory compliance frequency distribution - the majority of facilities [1] are either in full or substantial regulatory compliance.

An assumption within regulatory compliance is that full regulatory compliance, that is, 100 percent compliance with all rules [2], is the best (i.e., risk is minimized) possible scenario for the services being delivered and assessed. It is also assumed that all promulgated rules have an equal weight in their relative impact on the desired service delivery model, although this thinking has been changing over time regarding how rules are reviewed and complied with. This short treatise will examine the past 40 years of research delving into regulatory compliance measurement, and will provide some guidance to regulatory researchers and policy-makers as they move forward with both research and policy development related to rules. The data from these research studies have led to a Theory of Regulatory Compliance that demonstrates that substantial regulatory compliance - and not full regulatory compliance - is a more effective and efficient public policy as it relates to decision making on monitoring and licensing.

The results reported herein are drawn from human services delivery systems in the United States and Canada, such as early care and education, as well as child and adult residential services. The results are from state and provincial level licensing systems involving over 10,000 facilities serving over 100,000 clients. All the data are part of an international regulatory compliance database (https://data.mendeley.com/datasets/kzk6xssx4d/1) maintained at the Research Institute for Key Indicators and the Pennsylvania State University.

2. Methods

Alternate methodologies, logic models, and algorithms were developed directly from the Theory of Regulatory Compliance once it was determined that substantial regulatory compliance produced better results than full regulatory compliance. These methodologies created a differential monitoring or targeted monitoring approach based on risk assessment, which measures client morbidity and/or mortality when individual rule

---

*Corresponding author: Richard J. Fiene, Email:
rjf8@psu.edu, Phone: 717-598-8908, ORCID iD: http://ORCID: 0000-0001-6095-5085.

[1]The term "facilities" is used when referring to programs and/or facilities.

[2]The term "rules" is used when referring to rules and/or regulations.

non-compliance is assessed, and the determination of key statistical predictors for overall regulatory compliance [3].

Briefly, the above methodologies provide cost-effective and efficient means for the ongoing monitoring of human service delivery systems by selecting and reviewing only those rules that either have a positive impact on clients, statistically predict overall regulatory compliance, or protect the health and safety of clients [3]. Based on regulatory compliance historical data, decisions could be made as to the frequency and depth of the reviews or inspections. Abbreviated reviews (inspections in which a subset of rules are measured), such as licensing key indicator rules or risk assessment rules, would only be done in those facilities having a history of high regulatory compliance. Those facilities with a history of high regulatory non-compliance would continue to receive full regulatory compliance reviews as they did in the past.

3. Results

Prior to 1979, it was always assumed that there was a linear relationship between regulatory compliance measures and program quality measures of human service facilities. In a study conducted in that year, which compared results from early care and education programs, in particular child care centers, this assumption did hold up when one went from low regulatory compliance to substantial regulatory compliance. However, the results from substantial regulatory compliance to full (100 percent) regulatory compliance did not show the same linear relationship. Rather, it showed that those programs that were in substantial instead of full compliance were actually scoring higher on the program quality measures.

Since 1979, this result has been replicated in many other early care and education delivery system studies, both nationally in the United States (Head Start) [1] and in several states (Georgia, Indiana, Pennsylvania) [2]. In all these studies, one finds a non-linear - rather than a linear - relationship between regulatory compliance and the overall quality of the facilities being assessed.

4. Discussion

Based on the results above, there are several assumptions within regulatory compliance that need to be reconsidered:

1. Public policies that require full (100 percent) compliance with all rules may not be in the best interest of the clients being served, nor an effective use of limited regulatory resources. Potentially, emphasis on substantial regulatory compliance may be a more effective and efficient public policy related to client outcomes when it comes to their health, safety, and quality of life. Note that substantial compliance is still very high regulatory compliance (99-97 percent compliance with all rules) and produces positive client outcomes. As stated above, regulatory compliance data are extremely skewed and not normally distributed. There is very little variance in the data and the

majority of programs are in either full or substantial regulatory compliance.

2. If a jurisdiction focuses on a substantial regulatory compliance public policy it opens up many system enhancements, such as differential or targeted monitoring, risk assessment analysis, and statistical key indicator rules that have been demonstrated to be cost effective and efficient approaches to reviewing program performance. In a full regulatory compliance public policy approach, none of these system enhancements can be employed, with the possible exception of the key indicator approach as delineated in number four below.

3. If a jurisdiction takes the position that all rules are not equal, then a risk assessment or weighting approach becomes an alternative based on the assumption that certain rules place clients at greater risk of death, serious injury, or other types of harm.

4. Even if a jurisdiction does not have a licensing law that allows issuing licenses on the basis of substantial compliance, there is the possibility that key indicators could still be used for abbreviated reviews or inspections, if there is no prohibition in statute or regulation that expressly forbids the use of this approach, since key indicators statistically predict full regulatory compliance. In other words, all rules are statistically predicted to be in regulatory compliance based on the results of the key indicators. Therefore, technically, all rules have been reviewed albeit short of a full review or inspection.

5. Based on previous research, utilizing a risk assessment approach along with a key indicator approach is the most cost effective and efficient differential monitoring system model. The reason is that both predictive rules and those rules that place clients at greatest risk are always assessed when a site visit review or inspection is done. Many more jurisdictions use a risk assessment approach at this point, but there is a loss of predictive regulatory compliance by just using it.

6. Based on previous regulatory compliance history, only those facilities in high regulatory compliance would be eligible for abbreviated key indicator and risk assessment reviews, whereas those with a history of high regulatory non-compliance would continue to receive full regulatory compliance reviews. This gets at the essence of the differential monitoring approach, which is cost neutral. Regulatory resources may then be re-allocated from the abbreviated reviews to more in-depth full regulatory compliance reviews.

7. Based on the use of the key indicator and risk assessment methodologies within a differential monitoring approach, it is possible to identify over multiple jurisdictions if there are generic rules that meet the criteria of risk abatement and prediction. Such an application has occurred in the United States with the creation of early care

and education standards entitled *Caring for Our Children Basics*, published by the Administration for Children and Families, US Department of Health and Human Services (2015).

5. Conclusion

Regulatory compliance is relatively new in applying empirical evidence and basic scientific principles to its decision making. In the past, it had been dominated by case studies and long narrative reports that did not lend themselves to quantitative analysis. There is a need to more clearly apply empirical evidence and the scientific method to rule development. Certain assumptions, such as full regulatory compliance as a sound public policy, are lacking in empirical evidence. This treatise on a theory of regulatory compliance is provided for its heuristic value for both regulatory researchers and policymakers in rethinking some basic regulatory compliance assumptions. It is not about more or less, rules but finding the "right rules" that protect clients, predict overall regulatory compliance, and produce positive client outcomes.

6. Declaration of Conflicting Interest

The authors declare no conflicts of interest.

7. Article Information

This article was received March 14, 2019, in revised form April 11, 2019, and made available online May 9, 2019.

8. References

[1] Fiene, R. (2013). *Office of Head Start Key Indicator Project Report.* Retrieved from https://drfiene.files.wordpress.com/2018/04/rheadstart-report.pdf

[2] Fiene, R. (2014). *Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System.* Retrieved from https://drfiene.files.wordpress.com/2018/04/2014-georgia.pdf

[3] Fiene, R. (2016). *Early Childhood Program Quality Indicator and Improvement Model (ECPQIM) and Differential Monitoring Logic Model and Algorithm (DMLMA) Readings.* Retrieved from https://drfiene.files.wordpress.com/2018/03/1ecpqims-book-of-readings.pdf

Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System

Executive Summary

Richard Fiene, Ph.D.

The purpose of this study was to validate Georgia's process for determining if a state-regulated child care facility is compliant with basic state health and safety requirements. The process was developed by staff at Bright from the Start: Georgia Department of Early Care and Learning (DECAL). Currently Georgia utilizes a "Core Rule" risk assessment approach in which the health and safety rules deemed most crucial to ensure children's health and safety are used to compute a program's compliance status.

This validation study utilized a unique analytical model that compared licensing data with previous key indicator (*for readers not familiar with this term, please see the definitions on page 4 of the report*) research and ascertained if the Core Rules accurately indicated a program's overall compliance with the total population of licensing rules.

Additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that discerned between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Also licensing data were compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality. A Differential Monitoring Logic Model/Algorithm (*DMLMA©*) (Fiene, 2012) and a previous validation model (Zellman & Fiene, 2012) were used in the research.

One hundred and four child care centers (104 CCC) and 147 family child care (FCC) homes were assessed. Licensing data over a four-year period (2008-2012) and matching program quality data from a two-year period (2007-2008) were used in this study.

The study focused on three research questions:

1. Do the Core Rules CCCs and FCC homes serve as overall Key Indicators of compliance?
2. Does the Annual Compliance Determination Worksheet (ACDW) appropriately designate programs as compliant or non-compliant related to health and safety?
3. Are the Core Rules related to program quality?

The analysis demonstrated that the Core Rules did serve as key indicators, and these key indicators were identified for both center based and home based child care. The second analysis concluded that the ACDW computation did distinguish between compliant and non-compliant programs. Finally, the expected correlation between compliance and quality was found but only for state-funded Pre-K classrooms, not for family child care nor for preschool classrooms that were not part of the state-funded Pre-K.

Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System

Richard Fiene, Ph.D.

February 1, 2014

ABSTRACT

The purpose of this study was to validate Georgia's process for determining if a state-regulated child care facility is compliant with basic state health and safety requirements. The process was developed by staff at Bright from the Start: Georgia Department of Early Care and Learning (DECAL). Currently Georgia utilizes a "Core Rule" risk assessment approach in which the health and safety rules deemed most crucial to ensure children's health and safety are used to compute a program's compliance status. This validation study utilized a unique analytical model that compared licensing data with previous key indicator (*for readers not familiar with this term, please see the definitions on page 4 of the report*) research and ascertained if the Core Rules accurately indicated a program's overall compliance with the total population of licensing rules. Additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that discerned between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Also licensing data were compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality. A Differential Monitoring Logic Model/Algorithm (*DMLMA©*) (Fiene, 2012) and a previous validation model (Zellman & Fiene, 2012) were used in the research. Child care centers (CCC) and family child care (FCC) homes were assessed. The analysis demonstrated that the Core Rules did serve as key indicators, though this list should be reexamined. The second analysis concluded that the computation could be simplified. Finally, the expected correlation between compliance and quality was found but only in state-funded Pre-K classrooms; it was not found in preschool classrooms and could not be validated. Family child care could not be validated either. As a result of the study, recommendations were made to strengthen Georgia's system.

INTRODUCTION

Background of Georgia's Compliance Determination System

Similar to other states, Georgia has a licensing and monitoring system that oversees a diverse population of early care and learning programs across the state. The licensing and monitoring system of early care and learning programs is charged to Bright from the Start: Georgia Department of Early Care and Learning (DECAL), a state early education department that also oversees and administers Georgia's Pre-K Program, Child Care and Development Block Grant, the Child and Adult Care Food Program, and the Summer Food Service Program. In 2012, DECAL's licensing and monitoring system regulated approximately 6,300 early care and learning programs. The crux of this regulation is determining if the programs meet Georgia's health and safety rules. Programs that meet these rules are determined to be compliant.

In the mid 2000's, Georgia began experimenting with a process that determined whether or not a program was designated as compliant with the state's health and safety regulations by focusing on key Core Rules. These are health and safety rules deemed crucial to minimizing risk related to children's health and safety. Seventy-four rules out of the 456 that programs must follow were classified as Core Rules[1]. Core Rules are cited by severity (low, medium, high, extreme). It is important to note that this entails a risk assessment theoretical approach rather than a Key Indicator statistical approach. This means that the Core Rules were determined by content analysis rather than by a statistical procedure.

Though this system has undergone some slight revisions, this basic methodology is still in place:
1. All programs receive at least one full licensing study and one monitoring visit. At the licensing study all applicable rules are examined. At the monitoring visit, only Core Rules (or any rule that was not met at the licensing study) are examined.
2. If additional visits are conducted, the Core Rules are examined again at that time.
3. At the end of the fiscal year (June 30), each program receives a compliance determination. This determination is based on all visits (licensing study, monitoring visit, and other reviews). A standardized worksheet, Annual Compliance Determination Worksheet (ACDW), is used to make the computation that determines the designation.
4. The compliance status remains until the next determination one year later. Programs do not have an opportunity to contest the compliance determination, though programs have numerous opportunities to contest any citation.
5. At the conclusion of Fiscal Year 2012, approximately 91% of the programs were classified as compliant. A program's eligibility for certain services, acceptance into Quality Rated and Georgia's Pre-K Program, is impacted by the program's compliance determination.

Background of this Study

Since the compliance determination system has been used for several years, key policymakers at DECAL requested an external review to validate if the system was operating as intended. Are the Core Rules a sufficient subsample to measure a program's overall regulation with the state's health and safety regulations? Furthermore, does the compliance determination formula appropriately differentiate compliant programs from non-compliant programs? In other words, is the computation a viable way to make this designation? And finally, does compliance determination serve as a sufficient indicator for other aspects of quality not addressed in Georgia's health and safety rules?

The purpose of this study was to validate the aforementioned compliance determination process. This validation process utilized a unique analytical model that compared licensing data with previous key indicator research and ascertained if the Core Rules are an indication of a program's overall compliance with the total population of licensing rules. Second, additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that differentiated between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Finally, licensing data were

---

[1] The number of Core Rules was expanded in 2012 to include increased enforcement and sanctions regarding transportation. The new Core Rules were not part of this analysis.

compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality (see a further explanation of the sample in the Limitations Section of this report).

Specifically, the study addressed the following research questions:

1. **Do the Core Rules for child care centers (CCC) and family child care (FCC) homes serve as overall Key Indicators of compliance?**
2. **Does the Annual Compliance Determination Worksheet (ACDW) appropriately designate programs as compliant or non-compliant related to health and safety?**
3. **Are the Core Rules related to program quality?**

The following definitions are used in the study:

Core Rules = the rules determined to be of greatest importance and place children at greatest risk if not complied with. This approach is defined in the licensing literature as a risk assessment approach. Core Rules cover 12 regulatory areas and 74 specific rules. The Core Rules were the focal point of this validation study and are addressed in the first approach to validation – Standards and the first research question.

ACDW = Annual Compliance Determination Worksheet, the compliance decision-making system based on the Core Rules that can be used to determine the number of visits made to programs. The ACDW was the secondary focal point of this validation study and is addressed in the second approach to validation – Measures and the second research question.

Key Indicators = a differential monitoring approach that uses only those rules that statistically predict overall compliance with all the rules. In other words, if a program is 100% in compliance with the Key Indicators, the program will also be in substantial to full compliance with all rules. The reverse is also true in that if a program is not 100% in compliance with the Key Indicators, the program will also have other areas of non-compliance with all the rules. In this study, eight Key Indicators rules were identified for CCC and nine Key Indicators rules for FCC (See pages 15-16 for the specific indicators and additional detail about the methodology). These are in addition to the Core Rules.

Rule Violations or Citations = occurs when a program does not meet a specific rule and is cited as being out of compliance with that rule. These individual rule violations/citations are summed to come up with total violation/citation scores on the Core Rules and on the Licensing Studies.

Differential Monitoring = a relatively new approach to determining the number of licensing visits made to programs and to what rules are reviewed during these visits. Two measurement tools drive differential monitoring: one is a Weighted Risk Assessment, and the other is a Key Indicator checklist. Weighted Risk Assessments determine how often a program will be visited while Key Indicator checklists determine what rules will be reviewed in the program. Differential monitoring is a powerful approach when Risk Assessment is combined with Key Indicators because a program is reviewed by the most critical rules and the most predictive rules. See Figure 1 which presents a Logic Model & Algorithm for Differential Monitoring *(DMLMA©) (*Fiene, 2012).

Licensing Study = a comprehensive review of a program where all child care rules are reviewed.

Monitoring Visit = an abbreviated form of a visit and review in which only a select group (Core Rules) of child care rules are reviewed.

Program Quality = for the purposes of this study, quality was measured in child care centers by the *Early Childhood Environment Rating Scale-Revised (ECERS-R), Infant Toddler Environment Rating Scale-Revised (ITERS-R)* and in family child care homes by the *Family Child Care Environment Rating Scale-Revised (FCCERS-R).* The program quality measures were used as part of the third approach to validation – Outputs and the third research question.

Scoring for Licensing Variables/Data Collection Protocols:

Licensing Study = the total number of rule violations for a specific facility.

Core Rules = the total number of core rule violations.

ACDW/Compliance Designation = the annual compliance determination taken from the Annual Compliance Determination Worksheet. Compliant [C] was coded as "1" in the data base; Non-Compliant [NC] was coded as "0" in the data base.

Key Indicators = these were generated by a statistical methodology based upon the ability of the specific rule to predict full compliance with all the rules. Data from the Licensing Studies were used to make this determination of key indicator rule status.

METHODOLOGY AND ANALYTICAL FRAMEWORK

Licensing data over a four-year period (2008-2012) and matching program quality data from a two-year period (2007-2008) were used in this study. Specifically, data from 104 child care centers and 147 family child care homes were analyzed. Data from licensing studies (all rules) and monitoring visits (selected rules) were utilized. Program quality data were provided by researchers from the FPG Child Development Institute at the University of North Carolina at Chapel Hill (FPG), and the FPG research team matched these data points with the licensing data provided by DECAL (See the following website for the specific reports - *http://decal.ga.gov/BftS/ResearchStudyOfQuality.aspx)*. All the data were analyzed by the Research Institute for Key Indicators.

Two models were used to frame the analysis: a Validation Framework that uses four approaches (Zellman & Fiene, 2012) to validating quality rating and improvement systems (QRIS) being applied to licensing systems; and a *Differential Monitoring Logic Model and Algorithm (DMLMA©)(*Fiene, 2012) were employed to answer the three research questions for this Validation Study. The validation approaches are described below; the *DMLMA©* is described at the beginning of the Findings Section of this report.

The first validation approach deals with examining the validity of key underlying concepts by assessing if basic components and standards are the right ones by examining levels of empirical and expert support. For this study, this approach used Key Indicators to validate the Core Rules since Risk Assessment and Key Indicators are differential monitoring approaches. This answers the first research question.

The second validation approach deals with examining the measurement strategy and the psychometric properties of the measures used by assessing whether the verification process for each rule is yielding accurate results. Properties of the key rules can be measured through inter-rater reliability on observational measures, scoring of documentation, and inter-item correlations to determine if measures are psychometrically sound. Cut scores can be examined to determine the most appropriate ways to combine measures into summary ratings. For this study, the second validation approach validates the use of the ACDW and Core Rules by comparing compliance decisions with the Licensing Studies. This answers the second research question.

The third validation approach deals with assessing the outputs of the licensing process by examining the variation and patterns of program level ratings within and across program types to ensure that the ratings are functioning as intended. The approach examines the relationship of program level ratings to other more broadly based program quality measures and examines alternate cut points and rules to determine how well the ratings distinguish different levels of quality. For this study, this approach used data from Core Rules and Licensing Studies and data from earlier program quality studies (Maxwell, et al., 2009a,b; 2010) for validation. This answers the third research question.

Out of the four validation approaches (See Table 8), only three were utilized in this study. *The fourth validation approach* deals with how ratings are associated with children's outcomes. This approach examines the relationship

between program level ratings and selected child outcomes to determine whether higher program ratings are associated with better child outcomes. This approach did not have data that could be used in this study.

FINDINGS

The *DMLMA©* (See Figure 1) provides the conceptual model for assessing the overall effectiveness of Georgia's approach using Core Rules. In the model, the two main tools are Risk Assessment and Key Indicator measurements, which are created from a statistical analysis of the comprehensive licensing tool. The comprehensive licensing tool measures compliance with all rules. For the purposes of this study the Licensing Study represents the comprehensive licensing tool while the Core Rules represent a Risk Assessment tool. For the Program Quality tools, the ECERS-R, ITERS-R and FCCERS-R were utilized from an earlier program quality study by FPG Child Development Institute at the University of North Carolina at Chapel Hill (Maxwell, et al., 2009a,b; 2010). Georgia currently does not use a Key Indicator tool (see Table 1). With the DMLMA© analytical methodology, specific correlational thresholds are expected (please refer to Figure 1 on page 14).

TABLE 1

DMLMA© **Terminology**	**Georgia Examples and Data Sources**
Comprehensive Tool	Licensing Study
Program Quality Tool	ECERS-R and ITERS-R for CCC; FCCERS-R for FCC
Risk Assessment Tool	Core Rules
Key Indicators Tool	Not Present (Generated as part of this Study-see Tables 9/10)
Differential Monitoring Tool	ACDW Compliance Determination

Before presenting the findings for the validation approaches, some basic descriptive statistics are provided regarding the major variables in this study: Licensing Study, ACDW, Core Rules, and Key Indicators (see Table 2). The data are provided for both child care centers and family child care homes. It is clear from these basic descriptive statistics that the data distributions are very skewed in a positive fashion which means that there is very high compliance with all the major licensing variables for this study. In other words, the majority of programs are in substantial compliance with all the licensing rules and receive a compliant determination.

TABLE 2

Licensing Variable	**Mean**	**Range**	**SD**	**Skewness**	**Kurtosis**
Licensing Study (CCC)	5.51	25	5.26	1.47	2.11
ACDW (CCC)	0.75	1	0.44	-1.17	-0.64
Core Rules (CCC)	4.47	22	4.72	1.81	3.60
Key Indicators (CCC)	1.68	6	1.61	0.90	0.073
Licensing Study (FCC)	5.85	33	5.71	1.56	3.37
ACDW (FCC)	0.87	1	0.34	-2.23	3.03
Core Rules (FCC)	1.61	11	1.75	1.99	6.61
Key Indicators (FCC)	2.37	8	2.13	0.63	-0.57

Licensing Study Mean = the average number of total rule violations.
ACDW Mean = the average score for a determination of compliance (1) or non-compliance (0).
Core Rules Mean = the average number of core rule violations.
Key Indicators Mean = the average number of key indicator violations.

The findings are presented by the three validation approaches of Standards, Measures, and Outputs as well as the three research questions related to Key Indicators, Core Rules, and Program Quality.

1) **Validation of Standards (First Approach to Validation) for answering the first research question:** *Do the Core Rules for child care centers (CCC) and family child care (FCC) homes serve as overall key indicators of compliance?*

In this first approach to validation which focuses on Standards, Key Indicators were generated from the Licensing Studies because Core Rules (a Risk Assessment tool) and Key Indicators are both Differential Monitoring approaches (see Figure 1). The Core Rules were compared to the Key Indicators generated by the licensing data base and there was a .49 correlation for CCC (n = 104) and .57 correlation for FCC (n = 147) which indicates a

relationship between the Core Rules and Key Indicators at a p < .0001 significance level (Table 3). Also, the Key Indicators were correlated with the Licensing Study data and significant results were determined with r values of .78 (p < .0001) for CCC (n =104) and .87 (p < .0001) for FCC (n = 147). These results clearly met the expected *DMLMA©* thresholds between the key indicator rules with core rules (.50+) and licensing studies (.70+).

TABLE 3

Key Indicators with Core Rules and Licensing Study	r =	p <	n =
Key Indicators and Core Rules (CCC)	.49	.0001	104
Key Indicators and Licensing Study (CCC)	.78	.0001	104
Key Indicators and Core Rules (FCC)	.57	.0001	147
Key Indicators and Licensing Study (FCC)	.87	.0001	147

Table 3 begins to demonstrate how the Georgia Child Care Licensing system is utilizing the *DMLMA©* terminology from Table 1. With the generation of Key Indicators from this study, all the key elements within a differential monitoring system are present. This crosswalk to the *DMLMA©* will continue in Tables 4 & 5.

2) **Validation of Measures (Second Approach to Validation) for answering the second research question***: Is the Annual Compliance Determination Worksheet (ACDW) a valid measure in determining the overall health and safety compliance of Georgia's early care and learning programs?*

The Core Rules and the ACDW were compared to the Licensing Study data and compliance designation to determine the validation of the ACDW scoring protocol. There was a high correlation between the number of violations on the Core Rules and the total licensing violations on the Licensing Studies (r = .69; p < .0001)(Table 4). This result helps to validate that the ACDW is actually discriminating between high compliant and low compliant providers for CCC. For FCC, there was also a high correlation between the number of violations on the Core Rules and the total licensing violations on the Licensing Studies (r = .74; p < .0001). These results meet the *DMLMA©* thresholds of .50+ for Licensing Studies and Core Rules.

When Core Rules were correlated with the ACDW compliance decisions, there was a significantly high correlation for CCC (r = .76; p < .0001) and for FCC (r = .70; p < .0001). The key element of the ACDW scoring protocol is that the Core Rules distinguish between high and low compliant providers. The CCC/Core Rules and ACDW have been validated, as well as the FCC/Core Rules and ACDW because both the correlations were above the expected *DMLMA©* threshold (.50+).

TABLE 4

Core Rules with Licensing Studies and ACDW	r =	p <	n =
Core Rules and Licensing Studies (CCC)	.69	.0001	104
Core Rules and ACDW (CCC)	.76	.0001	104
Core Rules and Licensing Studies (FCC)	.74	.0001	147
Core Rules and ACDW (FCC)	.70	.0001	147

3) **Validation of Outputs (Third Approach to Validation) for answering the third research question:** *Are the Core Rules correlated with program quality?*

For this approach, programs were divided into those that had an ITERS-R score, an ECERS-R score for a preschool class, and an ECERS-R score for a Georgia's Pre-K class; and those that had only an ITERS-R score and an ECERS-R score for preschool. The sample was evenly divided. Since Georgia has placed substantial resources into its Pre-K program, it was thought that this analysis might suggest if there was anything different between programs with a Georgia's Pre-K class and those without.

When the Core Rules for CCC's were compared with program quality data (ECERS-R/PS + ITERS-R), a significant correlation was not found between CCC (r = .27) for programs with only preschool classrooms but was found for programs with Pre-K classrooms (ECERS-R/PK + ITERS-R) (r = .60). When Core Rules for FCC's were compared

to the FCC program quality data (FCCERS-R), the correlations were at a much lower level (r = .17) (See Table 5). However, these results are constrained by the limited range of the data; see the Limitation Section that follows this section.

Upon closer inspection of the correlations in Table 5 for CCC, it would appear that the CCC compliance system is more valid with the state-funded Pre-K programs (.48) than with the preschool programs (.21) because the correlations between the various Environment Rating Scales (ECERS-R + ITERS-R) are significant only when compared to the respective compliance with all rules on the Licensing Studies in the programs that have Pre-K programs. In making these comparisons, programs that had both ECERS-R and ITERS-R were combined and compared to the respective Licensing Study data (these data were reversed scored in which the number of violations were subtracted from a perfect score of 100). The differences are even more significant when you compare the Environment Rating Scales and the Core Rules where the Pre-K programs' correlation between the compliance with Core Rules and Environment Rating Scales is .60 and preschool programs is .27 while the FCC is .17.

Program quality data refer to data collected in earlier studies by researchers from FPG (Maxwell, et al., 2009a,b; 2010) in which FPG collected Environment Rating Scales (ECERS-R; ITERS-R; FCCERS-R) data on a representative sample of CCC and FCC (See (h*ttp://decal.ga.gov/BftS/ResearchStudyOfQuality.aspx*)**.** In comparing the program compliance and program quality data, the analyses supported the validation of the CCC for Pre-K only programs (*DMLMA©* threshold = .30+) but it was weaker for the FCC programs and not significant for preschool programs and therefore could not be validated. See Table 13 on page 17 for a further explanation of the CCC data distribution.

TABLE 5

Program Compliance and Quality Comparisons	r =	p <	n=
ECERS-R/PK + ITERS-R and Licensing Studies	.48	.001	45
ECERS-R/PK + ITERS-R and Core Rules	.60	.0001	45
ECERS-R/PS + ITERS-R and Licensing Studies	.21	ns	45
ECERS-R/PS + ITERS-Rand Core Rules	.27	ns	45
FCCERS-R and Licensing Studies	.19	.04	146
FCCERS-R and Core Rules	.17	.03	146

LIMITATION

The sampling for this study was based on previous studies (Maxwell, 2009a,b; 2010) completed by FPG in which program quality data were collected and analyzed. This study employed a subset of sites that were a representative sample of Georgia's child care licensing system. Not all of these sites could be used for this study because some had closed or some did not have the necessary data to make comparisons. So the sample at this point is one of convenience; however, 104 of the 173 CCC and 146 of the 155 FCC were used in this study, a significant number of the original representative sample. Also, when the Environment Rating Scales (ECERS-R, ITERS-R, FCCERS-R) scores were compared with the CCC and FCC samples, there were no significant differences (average difference was .01-.03) between the two study samples (See Table 6).

TABLE 6

Environment Rating Scale Scores	FPG	This Study
ECERS-R Pre-K Total Scale Scores	4.16	4.15
ECERS-R Preschool Total Scale Scores	3.39	3.42
ITERS-R Total Scale Scores	2.74	2.72
FCCERS-R Total Scale Scores	2.50	2.49

CONCLUSION

The CCC differential monitoring through the Core Rules/ACDW has been validated on the three approaches (Standards, Measures, and Outputs (Pre-K Program only)) and three research questions (Key Indicators, Core Rules, Program Quality (Programs with Georgia Pre-K only)) (See Table 7). The FCC differential monitoring through the Core Rules/ACDW was validated on the first validation approach (Standards) and first research question (Key Indicators); validated on the second validation approach (Measures) and second research question (Core Rules); but not validated on the third validation approach (Outputs) and third research question (Program Quality).

TABLE 7

Correlations

Validation Approach/Research Question	CCC Actual (Expected*)		FCC Actual (Expected)
1 STANDARDS/Key Indicators	VALIDATED		VALIDATED
Key Indicators x Core Rules	.49 (.50+)		.57 (.50+)
Key Indicators x Licensing Studies	.78 (.70+)		.87 (.70+)
2 MEASURES/Core Rules/ACDW[2]	VALIDATED		VALIDATED
Core Rules x Licensing Studies	.69 (.50+)		.74 (.50+)
Core Rules x ACDW	.76 (.50+)		.70 (.50+)
3 OUTPUTS/Program Quality	VALIDATED		NOT VALIDATED
Licensing Studies x ERS**/PK	.48 (.30+)	FCCERS	.19 (.30+)
Core Rules x ERS/PK	.60 (.30+)	FCCERS	.17 (.30+)
Licensing Studies x ERS/PS	------------		.21 (.30+)
Core Rules x ERS/PS	------------		.27 (.30+)

_____

DMLMA© Expected r Value Thresholds in Order to be Validated (Also see Figure 1 for additional details):
High correlations (.70+) = Licensing Studies x Key Indicators.
Moderate correlations (.50+) = Licensing Studies x Core Rules; Core Rules x ACDW; Core Rules x Key Indicators; Key Indicators x ACDW.
Lower correlations (.30+) = Program Quality Tools x Licensing Studies; Program Quality x Core Rules; Program Quality x Key Indicators.

Program Quality Tools = ECERS-R, ITERS-R, FCCERS-R.

**ERS = ECERS-R + ITERS-R
**PK = Pre-K program**
PS= Preschool program

A confounding of data occurred with the first two validation approaches because the Core Rules were influenced a great deal by the National Child Care Key Indicators (NCCKI) (Fiene, 2002) where 10 of the 13 Core Rules overlapped significantly with the NCCKI. This helped to increase the correlation between the Core Rules and the Licensing Studies because the Core Rules represented both risk assessment and key indicator rules. Using both risk assessment and key indicator rules together is an ideal differential monitoring approach (Fiene, 2012). Most states use one or the other but generally not together. By including the newly generated key indicators from this study where there is also overlap with the NCCKI, it should enhance the differential monitoring approach utilized by DECAL.

_____

[2.] ACDW decisions were compared with using severity as a factor and not using it as a factor in the scoring system with Core Rules. No significant differences were found between the two scoring systems; therefore, the results in this study represent Core Rule scores without severity included since this is the simpler model.

RECOMMENDATIONS

The following recommendations[3] can be made from this Licensing Differential Monitoring Validation Study.

1) **First research question/validation recommendation:** Revise the worksheet determination scoring relative to the visiting protocol by combining the Core Rules with a Key Indicator approach so that if any of the Core Rules or Key Indicators are out of compliance, then a full compliance review (Licensing Study) should be used. The present worksheet determination scoring protocol is overly complex. Just moving to a more comprehensive review (Licensing Study) based on non-compliance with the Core Rules will simplify the scoring protocol and make determinations more straightforward. If there is full (100%) compliance with the Core Rules and Key Indicators, then the next scheduled review of the program would be an abbreviated Monitoring Visit. If there is not 100% compliance with the Core Rules and Key Indicators, then the next scheduled review of the program would be a Licensing Study reviewing all child care rules. Based upon the compliance/non-compliance scores of the Licensing Study will determine how often the program will be visited. A revised Georgia Differential Monitoring System could potentially look like the following:
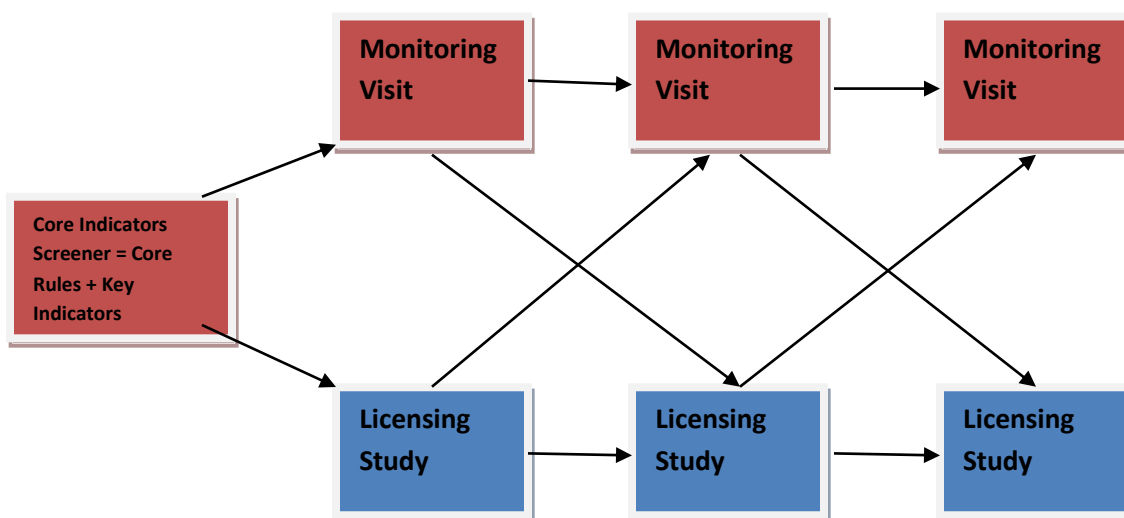
Compliance Decisions:
Core Indicators = Core Rules + Key Indicators – this becomes a screening tool to determine if a program receives a Licensing Study reviewing all child care rules or an abbreviated Monitoring visit continuing to review key indicator and core rules for their next visit.
Core Indicators (100%) = the next visit is a Monitoring Visit.. Every 3-4 years a full Licensing Study is conducted.
Core Indicators (not 100%) = The next visit is a Licensing Study where all rules are reviewed.
Compliance = 96%+ with all rules and 100% with Core Indicators. The next visit is a Monitoring Visit.
Non-compliance = less than 96% with all rules. The next visit is a Licensing Study..

2) **Second research question/validation recommendation:** Follow the development of weighted risk assessment tools as outlined by Fiene & Kroh (2000) in the *NARA Licensing Chapter* for CCC and FCC. It has been over 20 years since Core Rules were weighted. It is recommended that Core Rules be weighted every 10 years. Doing a weighted risk assessment would help confirm that the present Core Rules are the highest risk rules.

3) **Third research question/validation recommendation:** Confirm the CCC (ERS/PS) and FCC results by conducting a more recent program quality study that reflects all the changes made within the CCC and FCC systems. Although FCC program quality and Licensing Study and Core Rules reached statistical significance, the overall correlation was too low (Licensing Studies = .19; Core Rules = .17). With the CCC system the Pre-K program demonstrated significant correlations between ERS/PK and Licensing Study (.48) & Core Rules (.60) but not the Preschool program (ERS/PS: Licensing Studies = .21; Core Rules = .27).

_____

[3] These recommendations are drawn from the data in this study and previous studies conducted by the author in which the empirical evidence led to similar recommendations.

REFERENCES

Fiene (2012). *Differential monitoring logic model and algorithm (DMLMA©).* Middletown, PA: Research Institute for Key Indicators.

Fiene (2007). Child Development Program Evaluation & Caregiver Observation Scale, in T Halle (Ed.), *Early Care and Education Quality Measures Compendium*, Washington, D.C.: Child Trends.

Fiene (2003). Licensing related indicators of quality child care, *Child Care Bulletin*, Winter 2002-2003, pps 12-13.

Fiene (2002). *Thirteen indicators of quality child care: Research update*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services.

Fiene (1985). Measuring the effectiveness of regulations, *New England Journal of Human Services, 5*(2), 38-39.

Fiene (1981). A new tool for day care monitoring introduced by children's consortium, *Evaluation Practice, 1*(2), 10-11.

Fiene & Kroh (2000). Licensing Measurement and Systems, *NARA Licensing Curriculum*. Washington, D.C.: National Association for Regulatory Administration.

Fiene & Nixon (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14*(3), 198-214.

Griffin & Fiene (1995). *A systematic approach to policy planning and quality improvement for child care: A technical manual for state administrators*. Washington, D.C.: National Center for Clinical Infant Programs-Zero to Three.

Harms, T., Cryer, D.R., & Clifford, R.M. (2007). Family child care environment rating scale: Revised edition. New York: Teachers College Press.

Harms, T., Cryer, D., & Clifford, R. M. (2006). *Infant/toddler environment rating scale: Revised edition*. New York: Teachers College Press.

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale: Revised edition*. New York: Teachers College Press.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., Hume, K., & Crawford, G. (2009a). Georgia study of early care and education: Child care center findings. Chapel Hill, NC: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., Hume, K., & Crawford, G. (2009b). *Georgia study of early care and education: Findings from Georgia's Pre-K Program.* Chapel Hill, NC: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., & Hume, K., (2010). Georgia study of early care and education: Family child care findings. Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Morgan, Stevenson, Fiene, & Stephens (1986). Gaps and excesses in the regulation of child day care, *Reviews of Infectious Diseases--Infectious Diseases in Child Day Care: Management and Prevention, 8*(4), 634-643.

Zellman, G. L. and Fiene, R. (2012). *Validation of Quality Rating and Improvement Systems for Early Care and Education and School-Age Care,* Research-to-Policy, Research-to-Practice Brief OPRE 2012. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

TABLE 8 - FOUR APPROACHES TO VALIDATING A QRIS (Zellman & Fiene, 2012)

Approach	Activities and Purpose	Typical Questions Approach Addresses	Issues and Limitations
1. Examine the validity of key underlying concepts	Assess whether basic QRIS quality components and standards are the "right" ones by examining levels of empirical and expert support.	Do the quality components capture the key elements of quality? Is there sufficient empirical and expert support for including each standard?	Different QRISs may use different decision rules about what standards to include in the system.
2. Examine the measurement strategy and the psychometric properties of the measures used to assess quality	Examine whether the process used to document and verify each indicator is yielding accurate results. Examine properties of key quality measures, e.g., inter-rater reliability on observational measures, scoring of documentation, and inter-item correlations to determine if measures are psychometrically sound. Examine the relationships among the component measures to assess whether they are functioning as expected. Examine cut scores and combining rules to determine the most appropriate ways to combine measures of quality standards into summary ratings.	What is the reliability and accuracy of indicators assessed through program administrator self-report or by document review? What is the reliability and accuracy of indicators assessed through observation? Do quality measures perform as expected? (e.g., do subscales emerge as intended by the authors of the measures?) Do measures of similar standards relate more closely to each other than to other measures? Do measures relate to each other in ways consistent with theory? Do different cut scores produce better rating distributions (e.g., programs across all levels rather than programs at only one or two levels) or more meaningful distinctions among programs?	This validation activity is especially important given that some component measures were likely developed in low-stakes settings and have not been examined in the context of QRIS.

TABLE 8 (CONTINUED)

Approach	Activities and Purpose	Typical Questions Approach Addresses	Issues and Limitations
3. Assess the outputs of the rating process	Examine variation and patterns of program-level ratings within and across program types to ensure that the ratings are functioning as intended. Examine relationship of program-level ratings to other quality indicators to determine if ratings are assessing quality in expected ways. Examine alternate cut points and rules to determine how well the ratings distinguish different levels of quality.	Do programs with different program-level ratings differ in meaningful ways on alternative quality measures? Do rating distributions vary by program type, e.g., ratings of center-based programs compared to ratings of home-based programs? Are current cut scores and combining rules producing appropriate distributions across rating levels?	These validation activities depend on a reasonable level of confidence about the quality components, standards and indicators as well as the process used to designate ratings.
4. Examine how ratings are associated with children's outcomes.	Examine the relationship between program-level ratings and selected child outcomes to determine whether higher program ratings are associated with better child outcomes.	Do children who attend higher-rated programs have greater gains in skills than children who attend lower-quality programs?	Appropriate demographic and program level control variables must be included in analyses to account for selection factors. Studies could be done on child and program samples to save resources. Findings do not permit attribution of causality about QRIS participation but inferences can be made about how quality influences children's outcomes.

FIGURE 1- DIFFERENTIAL MONITORING LOGIC MODEL AND ALGORITHM (Fiene, 2012)
DMLMA© Applied to the Georgia Child Care Licensing System

CI + PQ => RA + KI => DM

Georgia Examples:

CI = Comprehensive Tool = Licensing Study (LS – All Rules)
PQ = Program Quality Tool = Environmental Rating Scales (ERS = ECERS-R, ITERS-R, FCCERS-R)
RA = Risk Assessment Tool = Core Rules (CR)
KI = Key Indicators Tool = presently Georgia does not have a KI
DM = Differential Monitoring Tool = ACDW (Compliance/Non-Compliance Decision)

A very important concept in this validation study is that the system employed by DECAL is a risk assessment approach rather than a key indicator methodology which is based upon predictor rules. The *DMLMA©* is a new methodology assessing the effectiveness and efficiency of Differential Monitoring systems being used by state regulatory agencies and provides the conceptual model for this study.

> *DMLMA© Thresholds:*
> *High Correlations (.70+) = CI x KI.*
> *Moderate Correlations (.50+) = CI x RA; RA x DM; RA x KI; KI x DM.*
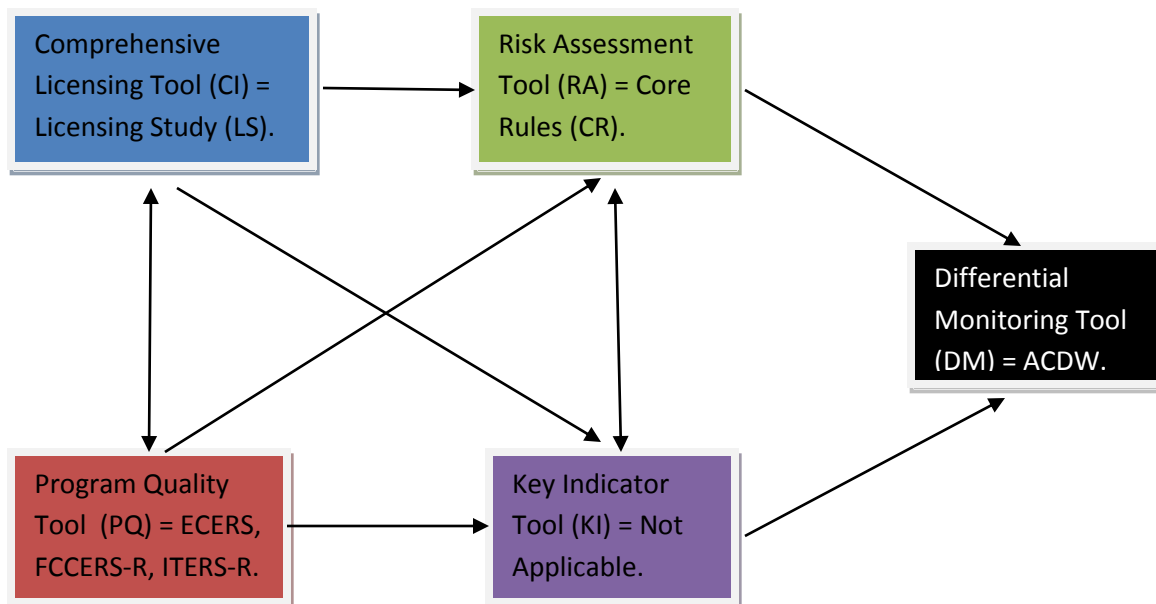> *Lower Correlations (.30+) = PQ x CI; PQ x RA; PQ x KI.*

Table 9 - Listing of Key Indicators for Georgia Child Care Centers with Phi Coefficients

591-1-1-25 (3) requires that the center and surrounding premises be clean, free of debris and in good repair. (Phi = .49)

591-1-1-.25 (13) requires that hazardous equipment, materials and supplies be inaccessible to children. (Phi = .46)

591-1-1-.26 (6) requires that outdoor equipment be free of hazards such as lead-based paint, sharp corners, rust and splinters. (Phi = .44)

591-1-1-.26 (8) requires the playground to be kept clean, free of litter and hazards. (Phi = .59)

591-1-1.26 (7) requires that a resilient surface be provided and maintained beneath the fall zone of climbing and swinging equipment. (Phi = .57)

591-1-1-.36 (6)(a-c) requires the center to maintain on the vehicle current information for each child including a) center and passenger information; b) emergency medical information and c) a passenger checklist. (Phi = .49)

591-1-1-.14 (1) requires that at least 50% of the caregiver staff have current first aid and CPR training. (Phi = .49)

591-1-1-.08 (a)-(f) requires the center to maintain a file for each child while such child is in care and for one year after that child is no longer enrolled…. (Phi = .44)

Table 10 - Listing of Key Indicators for Georgia Family Child Care Homes with Phi Coefficients

290.2.3-.11(2)(C) requires that fire drills be practiced monthly and shall be documented and kept on file for one year. (Phi = .51)

290-2-3-.11 (2)(f) requires that poisons, medicines, cleaning agents and other hazardous materials be in locked areas or inaccessible to children. (Phi = .61)

290-2-3-.11 (1)(f) requires the family day care home and any vehicle used to have a first aid kit….. (Phi = .57)

290-2-3-.07 (4) requires that the provider obtain ten clock hours of training in child care issues from an approved source within the first year and thereafter on an annual basis. (Phi = .58)

290-2-3-.08 (1)(a) requires the family day care home to maintain a file for each child that includes the child's name, birth date, parents or guardian's name, home and business addresses and telephone numbers. (Phi = .63)

290-2-3-.08 (1)(b) requires that the record for each child contain the names(s), address(es) and telephone number(s) of person(s) to contact in emergencies when the parent cannot be reached. (Phi = .57)

290-2-3-.08 (1)(b) requires the family day care home to maintain a file for each child that includes the name, address and telephone number of the child's physician to contact in emergencies. (Phi = .55)

290-2-3-.08 (1)(f) requires the family day care home to maintain a file for each child that includes known allergies, physical problems, mental health disorders, mental retardation or developmental disabilities which would limit the child's participation in the program. (Phi = .51)

290-2-3-.08 (1)(c) requires the family day care home to maintain a file for each child that includes evidence of age appropriate immunizations or a signed affidavit against such immunizations; enrollment in the home may not continue for more than 30 days without such evidence. (Phi = .72)

Table 11 - Key Indicator Formula Matrix for Generating Key Indicators*

	Providers In Compliance on Rule	Programs Out Of Compliance on Rule	Row Total
High Group**	A	B	Y
Low Group***	C	D	Z
Column Total	W	X	Grand Total

(* This computation occurred for each licensing rule)

****

Figure 2 - Key Indicator Statistical Methodology (Calculating the Phi Coefficient)

$$\phi \ = \ (A)(D) - (B)(C) \ \div \ \sqrt{(W)(X)(Y)(Z)}$$

_____

A = High Group + Programs in Compliance on Specific Rule.
B = High Group + Programs out of Compliance on Specific Rule.
C = Low Group + Programs in Compliance on Specific Rule.
D = Low Group + Programs out of Compliance on Specific Rule.
W = Total Number of Programs in Compliance on Specific Rule.
X = Total Number of Programs out of Compliance on Specific Rule.
Y = Total Number of Programs in High Group.
<u>*Z = Total Number of Programs in Low Group*</u>   .
**High Group = Top 25% of Programs in Compliance with all Rules.
***Low Group = Bottom 25% of Programs in Compliance with all Rules.

****

Table 12 – Phi Coefficient Decision Table

<u>Phi Coefficient Range</u>	<u>Characteristic of Indicator</u>	<u>Decision</u>
(+1.00) – (+.26)	Good Predictor	Include
(+.25) – (-.25)	Unpredictable	Do not Include
(-.26) – (-1.00)	Terrible Predictor	Do not Include

Table 13 - Comparison of the Pre-K and Preschool Programs

Compliance Level*	Pre-K ECERS-R**(N)	Preschool ECERS-R***(N)
100	4.88 (4)	3.40 (15)
99	4.13 (6)	4.35 (7)
98	4.38 (6)	3.89 (13)
97	3.99 (4)	3.15 (9)
96	4.36 (2)	3.16 (13)
95	4.60 (2)	3.53 (5)
90	3.43 (2)	2.56 (5)
80	2.56 (1)	2.38 (2)

*Compliance Level = the number of child care rule violations subtracted from 100.

100 = Full Compliance with Rules
99-98 = Substantial Compliance with Rules
97-90 = Medium Level of Compliance with Rules
80 = Low Level of Compliance with Rules

**Pre-K ECERS-R = average score of Pre-K Program classrooms as compared to the respective compliance levels. (N) = Sample Size.

***Preschool ECERS-R = average score of Preschool Program classrooms as compared to the respective compliance levels. (N) = Sample Size.

From this comparison there is more of a linear relationship between compliance levels and ECERS-R average scores for Pre-K Program classrooms than with the Preschool Program classrooms where there is more of a curvilinear or plateau effect at the upper end of compliance levels (Full Compliance). In order to attain the necessary correlational thresholds (+.30+) for validation for the third approach to validation, having a linear relationship rather than curvilinear will enhance this occurring. When a curvilinear or plateau effect occurs there is too great a likelihood that programs at a medium level of quality will be introduced into the highest (full) level of compliance. From a public policy standpoint this is an undesirable result.

The other item to note with the data distributions is that the Preschool ECERS-R data are more restricted than the Pre-K Program ECERS-R data. In other words, there is less variance in the Preschool Program ECERS-R data than in the Pre-K Program ECERS-R data.

There is an important limitation in these data that the reader must be aware of in not drawing any conclusions that the presence of a Pre-K Program classroom in any way is causing the change in licensing compliance. There is a relationship between the two but there is no assumption of causality.

Georgia Licensing Validation Technical Elements Appendix

Because of the nature of this report being a state's first attempt at fully validating it's Child Care Licensing Core Rule Differential Monitoring Approach utilizing the Zellman & Fiene (2012) Validation Framework and Fiene's DMLMA (2012) Model, certain questions surfaced regarding the terminology and the methodology being used in this report. This Technical Elements Appendix provides answers to specific questions that have been raised regarding these methodologies.

1. *How were the multiple years of data handled?*

The Licensing Study data used to make the comparisons are the facility reports that were the earliest facility observations so that these data would be closest to when the program quality data were collected. The other more recent Licensing Studies were not used in this comparison.

2. *If the Core Rules, Key Indicator, and Licensing Study values are counts of violations, how was the fact that different sites had different numbers of visits handled?*

Because only the earliest Licensing Study data was used, the number of visits were not an issue in the scoring.

3. *If the Core Rules, Key Indicator, and Licensing Study values are counts of violations, were all levels of violation risk (low, medium, high, extreme) handled the same?*

Yes, there were very few occurrences of high and extreme in the data base and also no significant differences were found when a sample of the rule violations with and without the levels of violation risk were compared. Therefore the simpler formula in which levels of violation risk were not used was selected.

4. *How did you determine the minimum correlations (DMLMA thresholds) for each analysis? Was this computed separately for this analysis or are the minimum correlations based on previous work?*

The DMLMA thresholds were determined from previous research work conducted by the author of this study on this model over the past 30 years. These were the average correlational thresholds that have been proposed for making validation determinations. The reason for utilizing the DMLMA model and thresholds is that the Zellman & Fiene (2012) Framework provides guidance in how to select specific validation approaches, what are the specific questions answered by the approach and what are the limitations of the particular approach. The DMLMA model builds upon this but provides a suggested scoring protocol by comparing correlational thresholds in a specific state to historical trends.

5. *Was Phi calculated for every rule in the licensing study? Can the full list be added to the appendix?*

Yes, Phi was calculated for every rule in the licensing study but most of them could not be computed because there was so few rule violations in the majority of the rules. This is typical of state licensing data sets and the full Phi comparisons are not depicted because it does not add any information to the state report.

6.   How did you determine which of the Licensing Study rules should be counted as Key Indicators?

The Key Indicator statistical methodology based upon a specific cut off point for the Phi Coefficient in which the p values were .0001 or less. This is a very stringent cut off point but it has been found historically that the p values needed to be lowered as the data distributions became more skewed with programs overall compliance levels increasing over time.

7.   How were sites that had no infant/toddler (i.e., no ITERS score) handled for the third validation approach? How were sites that had only a GA Pre-K (no preschool) handled?

For scoring purposes only those facilities that had both the ECERS and ITERS scores were used in making comparisons with the licensing data related to the third approach to validation. The GA Pre-K were scored and compared in the same way.

8.   On Table 13, why is the number of violation subtracted from 100 (rather than from the maximum possible)?

Generally this scoring is done because it is more intuitive to think in terms of 100% in compliance as a score of "100" rather than a score of "0".   This conversion is used in all state licensing reports that involve the DMLMA, Key Indicators and Risk Assessment Models.

FOR ADDITIONAL INFORMATION REGARDING THIS REPORT AND STUDY:

Bentley Ponder, Ph.D.
DECAL Director of Research and Evaluation
Bentley.Ponder@decal.ga.gov

Richard Fiene, Ph.D.
Research Institute for Key Indicators *LLC*
DrFiene@gmail.com

OFFICE OF HEAD START KEY INDICATOR PROJECT REPORT

Richard Fiene, Ph.D.

The purpose of this report is to present to the Office of Head Start (OHS) Key Indicators of their Head Start Performance Standards (HSPS) that have the ability to statistically predict substantial compliance with all Compliance Measures and ultimately the majority of HSPS's. The analytical and methodological basis of this approach is based upon a *Differential Monitoring Logic Model and Algorithm (DMLMA©)* (Fiene, 2012) (see Appendix 3). The DMLMA© is the 4th generation of an Early Childhood Program Quality Indicator Model (ECPQIM)(Fiene & Nixon, 1985; Griffin & Fiene, 1995; Fiene & Kroh, 2000). Only a portion of the *DMLMA©* model was utilized in this report which focused on key indicators, risk assessment, and program quality.

Definitions:

Risk Assessment (RA) - a differential monitoring approach that employs using only those rules, standards, or regulations that place children at greatest risk of mortality or morbidity if violations/citations occur with the specific rule, standard, or regulation.

Key Indicators (KI) - a differential monitoring approach that employs using only those rules, standards, or regulations that statistically predict overall compliance with all the rules, standards, or regulations. In other words, if a program is 100% in compliance with the Key Indicators the program will also be in substantial to full compliance with all rules, standards, or regulations. The reverse is also true in that if a program is not 100% in compliance with the Key Indicators the program will also have other areas of non-compliance with all the rules, standards, or regulations.

Differential Monitoring (DM) - this is a relatively new approach to determining the number of visits made to programs and what rules, standards, or regulations are reviewed during these visits. There are two measurement tools that drive differential monitoring, one is Weighted Risk Assessment tools and the other is Key Indicator checklists. Weighted Risk Assessments determine how often a program will be visited while Key Indicator checklists determine what rules, standards, or regulations will be reviewed in the program. Differential monitoring is a very powerful approach when Risk Assessment is combined with Key Indicators because a program is reviewed by the most critical rules, standards, or regulations and the most predictive rules, standards, or regulations. See Appendix 3 which presents a Logic Model & Algorithm for Differential Monitoring *(DMLMA©)*(Fiene, 2012).

Program Quality (PQ) - for the purposes of this study this was measured via the CLASS – Classroom Assessment Scoring System. The CLASS has three sub-scales (ES = Emotional Support, CO = Classroom Organization, and IS = Instructional Support). The CLASS is a tool that is identified in the research literature as measuring classroom quality similar to the ERS tools.

Early Childhood Program Quality Indicator Model (ECPQIM) – these are models that employ a key indicator or dashboard approach to program monitoring.   Major program monitoring systems in early care and education are integrated conceptually so that the overall early care and education system can be assessed and validated. With these models, it is possible to compare results obtained from licensing systems, quality rating and improvement systems (QRIS), risk assessment systems, key indicator systems, technical assistance, and child development/early learning outcome systems. The various approaches to validation are interposed within this model and the specific expected correlational thresholds that should be observed amongst the key elements of the model are suggested. **Key Elements** of the model are the following (see Appendix 3 for details): **CI** = state or federal standards, usually rules or regulations that measure health and safety - *Caring for Our Children or Head Start Performance* Standards will be applicable here. **PQ** = Quality Rating and Improvement Systems (QRIS) standards at the state level; *ERS (ECERS, ITERS, FDCRS), CLASS, or CDPES* (Fiene & Nixon, 1985). **RA** = risk assessment tools/systems in which only the most critical rules/standards are measured. *Stepping Stones* is an example of this approach**. KI** = key indicators in which only predictor rules/standards are measured. The *Thirteen Indicators of Quality Child Care* is an example of this approach. **DM** = differential monitoring decision making in which it is determined if a program is in compliance or not and the number of visits/the number of rules/standards are ascertained from a scoring protocol.   **PD** = technical assistance/training and/or professional development system which provides targeted assistance to the program based upon the **DM** results. **CO** = child outcomes which assesses how well the children are developing which is the ultimate goal of the system.

The organization of this report is as follows:

1) **The first section** will provide an overall analysis the Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) programs[1,4] ;

2) **The second section** will provide analyses of the various content areas (CA) within the HSPS[4];

3) **The third section** will provide analyses of the relationship between the HSPS as measured by compliance with the Compliance Measures (CM) and the program quality scores (CLASS scores)[3];

4) **The fourth and final section** will provide the analyses that produced the key indicators (KI) and recommendations in how it could be used.[2]

The source of data for this report is all the Tri-Annual On-Site Monitoring visits for 2012 which consisted of 422 reviews of programs across the country. There were 191 Head Start (HS) only programs, 33 Early Head Start (EHS) only programs, and 198 Head Start/Early Head Start (HS/EHS) programs reviewed. This is a representative sample of Head Start and Early Head Start programs nationally representing approximately 25% of the total number of Head Start programs.

Before proceeding with the results of this study, a few clarifying and definitional terms need to be highlighted. In the *2012 edition of OHS On-Site Review Protocol* and the *2013 OHS Monitoring Protocol*, Compliance Indicators (CI) and Key Indicators (KI) are respectively mentioned. In the licensing literature, when the term "Indicators" is used it refers to standards/rules that are predictive of overall compliance with all rules/standards. However, as defined by OHS, indicators (CI/KI) are used within the context of risk assessment which means that these indicators are the standards which are most important/critical

to the OHS in their monitoring reviews. These indicators therefore are not predictive in essence. That is the focus of this report/study which is to determine which of these indicators are predictive of overall compliance with all the compliance/key indicators. This is a common misconception in the human service regulatory field where risk assessment tools and key indicator tools purposes are confused. As we move forward please keep the definitions in mind related to the distinctions and functionality of risk assessment and key indicators.

For the purposes of this study, 131 Compliance Measures (CM), organized into seven (7) Content Areas (CA), were reviewed and analyzed. The seven content areas are the following: Program Governance; Management Systems; Fiscal Integrity; Eligibility, Recruitment, Selection, Enrollment, and Attendance; Child Health and Safety; Family and Community Engagement; Child Development and Education. Ten CM's were from Program Governance (GOV), 10 were from Management Systems (SYS), 22 were from Fiscal Integrity (FIS), 11 were from Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA), 34 were from Child Health and Safety (CHS), 16 were from Family and Community Engagement (FCE), and 28 were from Child Development and Education (CDE)[4].

Section 1 - Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) programs

In order to determine if analyses needed to be performed separately on Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) combined programs, the first series of analyses were performed to determine if any statistically significant differences existed amongst these three groups. This is a very important first analysis because it will help to determine the stability of the sample selected and of the overall system. In other words, is there a good deal of consistency across all service types: HS, EHS, and HS/EHS.

Based upon Table 1, no statistically significant differences were determined amongst the three groups (HS, EHS, HS/EHS) with Compliance Measures (CM) or CLASS (ES, CO, IS) Scores indicating that using the full 422 sample and not having to do separate analyses for the three groups was the correct analytical framework. However, where it is appropriate, any statistically significant differences amongst the various program types will be highlighted.

Table 1 – Head Start, Early Head Start, & Head Start/Early Head Start With CM and CLASS/ES, CO, IS

Program Type	CM(N)	CLASS/ES(N)	CLASS/CO(N)	CLASS/IS(N)
Head Start (HS)	3.72(191)	5.88(186)	5.43(186)	2.97(186)
Early Head Start (EHS)	2.67(33)	-----*	-----*	-----*
Head Start (HS/EHS)	3.07(198)	5.91(198)	5.47(198)	3.00(198)
Totals	**3.33(422)**	**5.89(384)**	**5.45(384)**	**2.98(384)**
Statistical Significance	**NS**	**NS**	**NS**	**NS**

CM = Compliance Measures (Average Number of Violations)
CLASS/ES = CLASS Emotional Support Average Score
CLASS/CO = CLASS Classroom Organization Average Score
CLASS/IS = CLASS Instructional Support Average Score
NS = Not Significant
N = Number of Programs

*CLASS data were not collected in EHS.

The average number of violations with the Compliance Measures for Head Start (3.72), Early Head Start (2.67) and Head Start/EHS (3.07) was not significant in utilizing a One-Way ANOVA. There were 191 Head Start (HS) programs, 33 Early Head Start (EHS) programs, and 198 Head Start (HS/EHS) programs.

Comparisons were also made with Head Start and Head Start/EHS on the various CLASS sub-scales (ES = Emotional Support, CO = Classroom Organization, and IS = Instructional Support) and no significant differences were found between these two groups. The EHS (n = 33) was not used because CLASS data were not collected in these programs.

The practical implication of the above results is that the same monitoring tools and the resulting Head Start Key Indicator (HSKI) to be developed as a result of this study can be used in the three main types of programs: Head Start, Early Head Start, and Head Start/EHS. There is no need to have separate tools.

Section 2 - Content Areas

The second series of analyses was to look more closely at the 7 content areas (CA) to measure demographically any differences amongst the various areas. In order to do this a weighted average had to be determined in order to compare the various areas because of the differences in the number of Compliance Measures (CM) used in each content area. Table 2 provides the results of these analyses. For the total sample of 422 sites, Management Systems (SYS) Content Area (CA) had the highest number of violations with the Compliance Measures (CM) with 359. The SYS/CA also had the highest average number of violations with 35.90 because there were only 10 CM. For the total sample of 422 sites, the lowest number of violations was in the Family and Community Engagement (FCE) Content Area (CA) with 48 violations with CM. It also had the lowest average number of violations with 3.00.

For the Head Start only sites (n = 191), a similar distribution as with the total sample (n = 422) is depicted in which Management Systems (SYS) Content Area (CA) had the highest number of violations with the Compliance Measures (CM) with 192. The SYS/CA also had the highest average number of violations with 19.20 because again there were only 10 CM. The lowest number of violations was in the Family and Community Engagement (FCE) Content Area (CA) with 20 violations with CM. It also had the lowest average number of violations with 1.25.

For the Early Head Start only (n = 33) and the Head Start/Early Head Start (n = 198) sites, the ranking of the various Content Areas changed somewhat with the total number of violations and the average number of violations from the Total Sample (n = 422) and the Head Start only (n = 191) sites but not dramatically. For example, the Family and Community Engagement (FCE); Child Development and Education (CDE); and the Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA) Content Areas switched rankings in which it had the fewest total violations and the average number of violations (see Table 2).

Table 2 – Comparing Content Areas and Program Types

Content Areas	Total Violations/(Rank)				Average # of Violations/(Rank)				CM #
	TOT	HS	EHS	HS/EHS	TOT	HS	EHS	HS/EHS	
FCE	48(1)	20(1)	2(1)	26(2)	3.00(1)	1.25(1)	0.125(1)	1.63(2)	16
ERSEA	62(2)	37(2)	6(3)	19(1)	5.64(3)	3.36(3)	0.545(3)	1.73(3)	11
CDE	91(3)	43(3)	5(2)	43(3)	3.25(2)	1.54(2)	0.179(2)	1.54(1)	28
GOV	150(4)	94(4)	6(3)	50(4)	15.00(6)	9.40(6)	0.600(4)	5.00(5)	10
FIS	255(5)	114(5)	23(7)	118(5)	11.59(5)	5.18(5)	1.045(6)	5.36(6)	22
CHS	333(6)	151(6)	22(6)	160(7)	9.79(4)	4.44(4)	0.647(5)	4.71(4)	34
SYS	359(7)	192(7)	20(5)	147(6)	35.90(7)	19.20(7)	2.000(7)	14.70(7)	10

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS =CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

TOT = TOTAL NUMBER OF SITES, FULL SAMPLE OF 422 SITES
HS = HEAD START ONLY PROGRAMS
EHS = EARLY HEAD START ONLY PROGRAM
HS/EHS = HEAD START AND EARLY HEAD START COMBINED PROGRAMS
CM = NUMBER OF COMPLIANCE MEASURES

TOTAL VIOLATIONS = ALL THE VIOLATIONS FOR A SPECIFIC CONTENT AREA.
AVERAGE # OF VIOLATIONS = THE TOTAL VIOLATIONS FOR A SPECIFIC CA DIVIDED BY THE NUMBER OF COMPLIANCE MEASURES FOR THAT SPECIFIC CONTENT AREA.
RANK = HOW EACH CONTENT AREA COMPARES TO THE OTHER CONTENT AREAS FOR THE RESPECTIVE PROGRAM TYPE.

For the total sample (n = 422), other CA's had different configurations between the total number of violations and the average number of violations as demonstrated by CHS – Child Health and Safety in which there was a total of 333 violations but the average number of violations was 9.79 because there were 34 Compliance Measures (CM). Program Governance (GOV) had 150 total violations and a weighted-average of 15 violations with 10 CM. Child Development and Education (CDE) had 91 total violations and a weighted-average of 3.25 violations. Fiscal Integrity (FIS) had 255 total violations and a weighted-average of 11.59 violations. And lastly, Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA) had 62 total violations and a weighted-average of 5.64 violations.

The Head Start only (HS = 191), Early Head Start only (EHS = 33), and the Head Start/Early Head Start (HS/EHS = 198) programs followed a similar pattern as with the total sample (n = 422).   This indicates a great deal of consistency in the sample drawn. See Appendix 4 for violation data for all 131 Compliance Measures.

The practical implication of the above findings is that certain Content Areas (SYS, GOV, FIS) may need additional exploration by OHS because of their high rates of non-compliance with the Compliance Measures.

Section 3 – Program Quality

This section provides comparisons between the Compliance Measures (CM) data and the CLASS (ES, CO, IS) data. This is a very important section because there is always the concern that compliance with the HSPS has no relationship to program quality as measured by the CLASS. In Table 3, correlations were run between the CM data and the CLASS scores for Emotional Support (ES), Classroom Organization (CO), and Instruction Support (IS) for the Head Start only and the Head Start/Early Head Start programs. The EHS only programs were not included because CLASS data are not collected on these programs. The results are very positive and statistically significant in most cases. It is also important to note the very positive correlation between the Head Start Key Indicators (HSKI[2]) and CLASS. This result supports using the HSKI in monitoring Head Start.

Table 3 – Relationship Between Compliance Measures (CM), KI, and CLASS (ES, CO, IS) Scores

CLASS	Compliance Measures Content Areas								Key Indicators
	CM	FCE	ERSEA	CDE	GOV	FIS	CHS	SYS	KI
CLASS/ES	.22**	.13*	.15**	.15**	.11*	.05	.23**	.17**	.27**
CLASS/CO	.19**	.13*	.11*	.16**	.04	.06	.21**	.15**	.25**
CLASS/IS	.20**	.10	.12*	.12*	.13*	.06	.18**	.11*	.17**

CM Violations = Total Compliance Measure Violations

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS =CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

CLASS/IS = Average CLASS IS (Instructional Support) Score
CLASS/ES = Average CLASS ES (Emotional Support) Score
CLASS/CO = Average CLASS CO (Classroom Organization) Score

KI = Key Indicators Total Score

** p < .01
* p < .05

See Appendix 6 & 6A for the inter-correlations amongst all the Content Areas, HSKI, and Total Compliance with Compliance Measures.

These results are very important but it is equally important to look more specifically at the distribution of the Compliance Measures (CM) scores and their relationship to the CLASS data (see Appendix 5 for detailed graphic distributions and Appendix 6 & 6A for the inter-correlations amongst all the CA). When this is done a very interesting trend appears (see Table 3a) in which a definite plateau occurs as the scores move from more violations or lower compliance with the Compliance Measures (25-20 to 3-8 CM Violations) to fewer violations or substantial compliance with the Compliance Measures (1-2 CM Violations) and full compliance with the Compliance Measures (Zero (0) CM Violations).

Table 3a – Aggregate Scores Comparing CM Violations with CLASS Scores

CM Violations		IS	ES	CO	Number/Percent
0	(Full Compliance)	3.03	5.99	5.59	75/19%
1-2	(Substantial Compliance)	3.15	5.93	5.50	135/35%
3-8	(Mid-Compliance)	2.87	5.85	5.37	143/40%
9-19	(Lower Compliance)	2.65	5.71	5.32	28/6%
20-25	(Lowest Compliance)	2.56	5.52	4.93	3/1%
Significance		**F = 4.92; p < .001**	**F = 4.918; p < .001**	**F = 4.174; p < .003**	

CM Violations = Compliance Measure Violations (lower score = higher compliance)(higher score = lower compliance)
IS = Average CLASS IS (Instructional Support) Score
ES = Average CLASS ES (Emotional Support) Score
CO = Average CLASS CO (Classroom Organization) Score
#/% = Number of programs and Percent of programs at each level of compliance

When comparing these groupings in Table 3a the results from a One Way ANOVA were significant (F = 4.92; p < .001) for the CLASS/IS Scores. The average CLASS/IS Score when there were no CM Violations was 3.03. The average CLASS/IS Score when there were 1-2 CM Violations was 3.15. The average CLASS/IS Score when there were 3-8 CM Violations was 2.87. The average CLASS/IS Score when there were 9-19 CM Violations was 2.65. And finally, the average CLASS/IS Score when there were 20-25 violations was 2.56. The results were very similar with the CLASS/ES and CLASS/CO scores as well in which the results from a One Way ANOVA were statistically significant for the CLASS/ES (F = 4.918; p < .001) and for the CLASS/CO (F = 4.174; p < .003). These results clearly demonstrate that being in full or substantial compliance with the Compliance Measures correlates with more positive scores on the CLASS. Approximately 55% of the Head Start programs are at the full or substantial compliance level.

The practical implication of the above findings is that placing equal emphasis on full as well as substantial compliance with the Compliance Measures could be an acceptable public policy decision.

Section 4 – Head Start Key Indicators (HSKI)

The fourth and final section of this report is in some ways the most important since this is the focus of the study: developing statistically predictive Key Indicator (KI) Compliance Measures (CM) – the Head Start Key Indicators (HSKI).

These are the statistically predictive Key Indicators based upon the KI methodology, correlations with the CLASS/ES, CO, IS, and correlations with the CM Total Violation scores. Table 4 lists the results while Appendix 1 has the specific KI's content specified. Appendix 2 depicts the KI Formula Matrix. Only those Compliance Measures (CM) that had significant results on three of the five correlations were selected to be Head Start Key Indicator Compliance Measures (HSKI).

The methodology used to generate the Compliance Measure Key Indicators sorted the top 20% of programs in compliance and compared this group to the bottom 27% of programs in compliance. The middle 53% of programs were not used in order to determine the Key Indicators. These cut off points

were determined by the compliance distribution in which 20% of the programs were in 100% compliance while 27% of the programs had compliance scores of 95% or less.

Table 4 – Head Start Key Indicator (HSKI) Compliance Measures (CM) and CLASS and Total Violations

HSKI/CM (2013)	Phi	CLASS/ES	CLASS/CO	CLASS/IS	Total Violations
CDE4.1	.28***	.10*	ns	ns	.30***
CHS1.1	.39***	.15**	.16**	ns	.39***
CHS1.2	.33***	.18**	.15**	.10*	.36***
CHS2.1	.49***	.18**	.15**	ns	.54***
CHS3.10	.39***	.11*	.11*	ns	.24***
GOV2.1	.31***	.11*	ns	ns	.46***
SYS2.1	.47***	.15**	.16**	.14**	.55***
SYS3.4	.58***	.13*	.10*	ns	.36***

Phi = the phi coefficient which statistically predicts compliance with the full set of CM's.

CLASS/ES = correlations between the specific CM and this specific scale of the CLASS.
CLASS/CO = correlations between the specific CM and this specific scale of the CLASS.
CLASS/IS = correlations between the specific CM and this specific scale of the CLASS.

Total Violations = correlations between the specific CM and the total number of CM violations for each program.

*   p < .05
**   p < .01
****   p < .001*
ns = not significant

Separate Key Indicators were run for just Head Start only and Head Start/Early Head Start programs but the key indicators were only a subset of the above list, albeit a shorter list in each case. Based upon those phi coefficients, it was determined that using the above list for all Head Start only, Early Head Start, and Head Start/Early Head Start was a more efficient and effective way to monitor all the programs with one list of indicators rather than having separate key indicators for program types. The separate phi coefficients run for Head Start only and Head Start/Early Head Start programs did not show any significant differences because they were sub-samples of the overall sample drawn.

Section 4A – Suggested Use of the HSKI for Head Start Program Monitoring

Now that Key Indicators have been generated, the next question is how to use HSKI in the program monitoring of Head Start. A possible way in which the HSKI could be used would be the following (see Figure 1) in which a differential monitoring approach could be used:

All programs would be administered the HSKI. If there is full (100%) compliance with the Head Start Key Indicators (HSKI) then the next scheduled review of the program would be an Abbreviated Monitoring Visit (AMV). If there is not 100% compliance with the Head Start Key Indicators (HSKI) then the next scheduled review of the program would be a Full Monitoring Visit (FMV) in which all Compliance Measures are reviewed. Based upon the results of the FMV a determination could be made regarding a compliance or non-compliance decision (see Figure 1) and how often the program will be visited.

Figure 1 – Head Start Key Indicator (HSKI) Compliance Measures Differential Monitoring Model
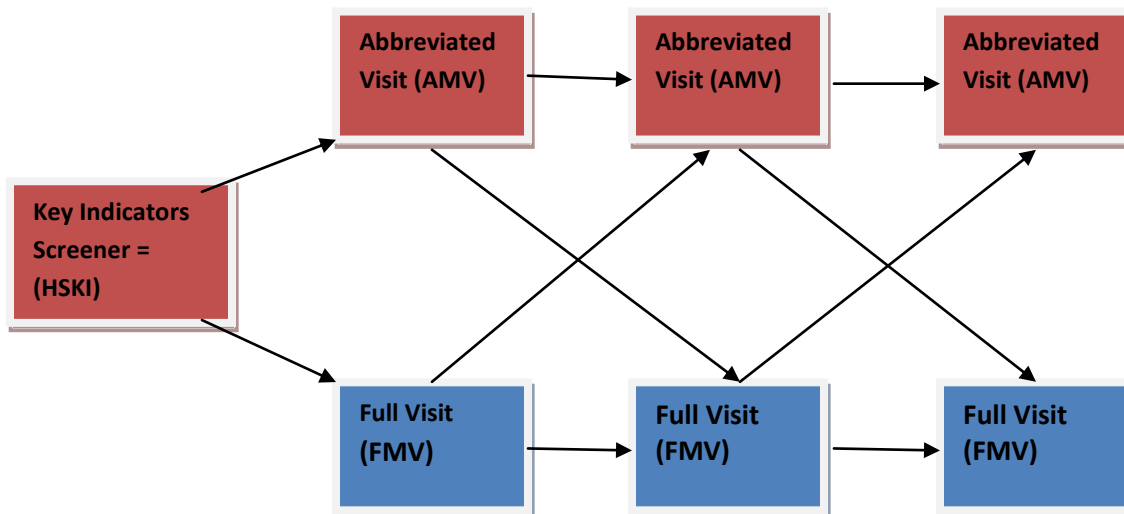
Compliance Decisions:

Head Start Key Indicators (HSKI) – this becomes a screening tool to determine if a program receives an AMV OR FMV visit.

HSKI (100%) = *For the next visit, an Abbreviated Monitoring Visit (AMV) is conducted. Every 3-4 yrs a full Monitoring is conducted.*

HSKI (not 100%) = *For the next visit, a Full Monitoring Visit (FMV) is conducted and all CMs are reviewed.*

Compliance = *98%+ with all CMs which indicates substantial to full compliance and 100% with HSKI. For the next visit, an Abbreviated Monitoring Visit (AMV) is conducted.*

Non-compliance = *less than 98% with all CMs which indicates low compliance. For the next visit a Full Monitoring Visit (FMV) is conducted.*

Moving to a differential monitoring system could provide a cost effective and efficient model for Head Start program monitoring. This revision to the Head Start program monitoring system would combine a risk assessment and key indicator approach (see Appendix 3) in determining what compliance measures to review, how often, and how comprehensive a review should be utilized. It would continue to focus on the most critical compliance measures that statistically predict overall compliance with the full complement of compliance measures.

See Appendix 7 – Figure 2 for how the above differential monitoring system could impact the present Head Start Tri-Annual Review Monitoring System. In this appendix, a cost neutral monitoring system is proposed based upon the above DMLMA/Key Indicator Model.

References

Fiene (2013). *Differential monitoring logic model & algorithm*, Research Institute for Key Indicators, Middletown, Pennsylvania.

Fiene (2003). Licensing related indicators of quality child care, *Child Care Bulletin*, Winter 2002-2003, pps 12-13.

Fiene (1985). Measuring the effectiveness of regulations, *New England Journal of Human Services, 5*(2), 38-39.

Fiene (1981). A new tool for day care monitoring introduced by children's consortium, *Evaluation Practice, 1*(2), 10-11.

Fiene & Kroh (2000). Licensing Measurement and Systems, *NARA Licensing Curriculum*. Washington, D.C.: National Association for Regulatory Administration.

Fiene & Nixon (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14*(3), 198-214.

Griffin & Fiene (1995). *A systematic approach to policy planning and quality improvement for child care: A technical manual for state administrators*. Washington, D.C.: National Center for Clinical Infant Programs-Zero to Three.

Footnotes

1) PIR Dashboard Key Indicators could not be generated because the PIR data demonstrated little statistical predictive ability to be useful for discriminating between high and low compliant programs or program quality with the exception of staff having CDA's.

2) The correlation between Compliance Measures (CM) and the statistically predictive Key Indicators (HSKI) was .77 which exceeds the expected correlation threshold.

3) The correlations between the CLASS/ES, CO, IS and Key Indicators were the following: .27, .25, .17 respectively. The correlations between KI and ES and CO were higher than the correlations between CM and ES, CO as reported earlier in this report. The correlation between IS and CM was higher .20 than KI and IS (.17).

4) Because this study spans the 2012 Review Protocol and 2013 Monitoring Protocol, Compliance Indicators and Compliance Measures are used interchangeably with a preference given to using Compliance Measures (CM) in this report. There are 139 Compliance Indicators; 115 Compliance Measures, but for the purposes of this study 131 Compliance Measures were available in the 2012 Head Start data base drawn for this study.

For additional information regarding this report, please contact:
Richard Fiene, Ph.D., Director
Research Institute for Key Indicators
41 Grandview Drive
Middletown, PA. 17057
DrFiene@gmail.com
RIKI.Institute@gmail.com
717-944-5868 Phone and Fax
http://RIKInstitute.wikispaces.com
http://pennstate.academia.edu/RickFiene

February 2013 (revised March, April 2013)

Appendix 1 – Head Start Key Indicators (HSKI) Compliance Measures Content

CM	Content	Regulations/Law
CDE4.1*	The program hires teachers who have the required qualifications, training, and experience.	1304.52(f), 645A(h)(1), 648A(a)(3)(B)(i), 648A(a)(3)(B)(ii), 648A(a)(3)(B)(iii)
CHS1.1	The program engages parents in obtaining from a health care professional a determination of whether each child is up to date on a schedule of primary and preventive health care (including dental) and assists parents in bringing their children up to date when necessary and keeping their children up to date as required.	1304.20(a)(1)(ii), 1304.20(a)(1)(ii)(A), 1304.20(a)(1)(ii)(B)
CHS1.2	The program ensures that each child with a known, observable, or suspected health, oral health, or developmental problem receives follow-up and further testing, examination, and treatment from a licensed or certified health care professional.	1304.20(a)(1)(iii), 1304.20(a)(1)(iv), 1304.20(c)(3)(ii)
CHS2.1	The program, in collaboration with each child's parent, performs or obtains the required linguistically and age-appropriate screenings to identify concerns regarding children within 45 calendar days of entry into the program, obtains guidance on how to use the screening results, and uses multiple sources of information to make appropriate referrals.	1304.20(a)(2), 1304.20(b)(1), 1304.20(b)(2), 1304.20(b)(3)
CHS3.10	Maintenance, repair, safety of facility and equipment	1304.53(a)(7)
GOV2.1*	Members of the governing body and the Policy Council receive appropriate training and technical assistance to ensure that members understand information they receive and can provide effective oversight of, make appropriate decisions for, and participate in programs of the Head Start agency.	642(d)(3)
SYS2.1	The program established and regularly implements a process of ongoing monitoring of its operations and services, including delegate agencies, in order to ensure compliance with Federal regulations, adherence to its own program procedures, and progress towards the goals developed through its Self-Assessment process.	1304.51(i)(2), 641A(g)(3)
SYS3.4	Prior to employing an individual, the program obtains a: Federal, State, or Tribal criminal record check covering all jurisdictions where the program provides Head Start services to children; Federal, State, or Tribal criminal record check as required by the law of the jurisdiction where the program provides Head Start services; Criminal record check as otherwise required by Federal law	648A(g)(3)(A), 648A(g)(3)(B), 648A(g)(3)(C)

*** FY 2013 Office of Head Start Monitoring Protocol (October 26, 2013) Compliance Measures**

Appendix 2: Key Indicator Formula Matrix for HSKI – Head Start Key Indicators

	Providers In Compliance	*Programs Out Of Compliance*	*Row Total*
High Group	*A*	*B*	*Y*
Low Group	*C*	*D*	*Z*
Column Total	*W*	*X*	**Grand Total**

Key Indicator Statistical Methodology (Calculating the Phi Coefficient):

$$\phi \;=\; (A)(D) - (B)(C) \;\div\; \sqrt{(W)(X)(Y)(Z)}$$

_____

A = High Group + Programs in Compliance on Specific Compliance Measure.
B = High Group + Programs out of Compliance on Specific Compliance Measure.
C = Low Group + Programs in Compliance on Specific Compliance Measure.
D = Low Group + Programs out of Compliance on Specific Compliance Measure.

W = Total Number of Programs in Compliance on Specific Compliance Measure.
X = Total Number of Programs out of Compliance on Specific Compliance Measure.
Y = Total Number of Programs in High Group.
Z = Total Number of Programs in Low Group.

_____

High Group = Top 20% of Programs in Compliance with all Compliance Measures.
Low Group = Bottom 27% of Programs in Compliance with all Compliance Measures.

Phi Coefficient Range	Characteristic of Indicator	Decision
(+1.00) – (+.26)	Good Predictor	Include on HSKI
(+.25) – (0)	Too Easy	Do not Include
(0) – (-.25)	Too Difficult	Do not Include
(-.26) – (-1.00)	Terrible Predictor	Do not Include

Appendix 3

DIFFERENTIAL MONITORING LOGIC MODEL AND ALGORITHM (Fiene, 2012) *DMLMA©* Applied to the Office of Head Start Program Monitoring Compliance System

CI + PQ => RA + KI => DM

Head Start Examples:

CI = Head Start Performance Standards (HSPS)
PQ = CLASS ES, IS, CO (CLASS)
RA = Compliance Measures (CM)
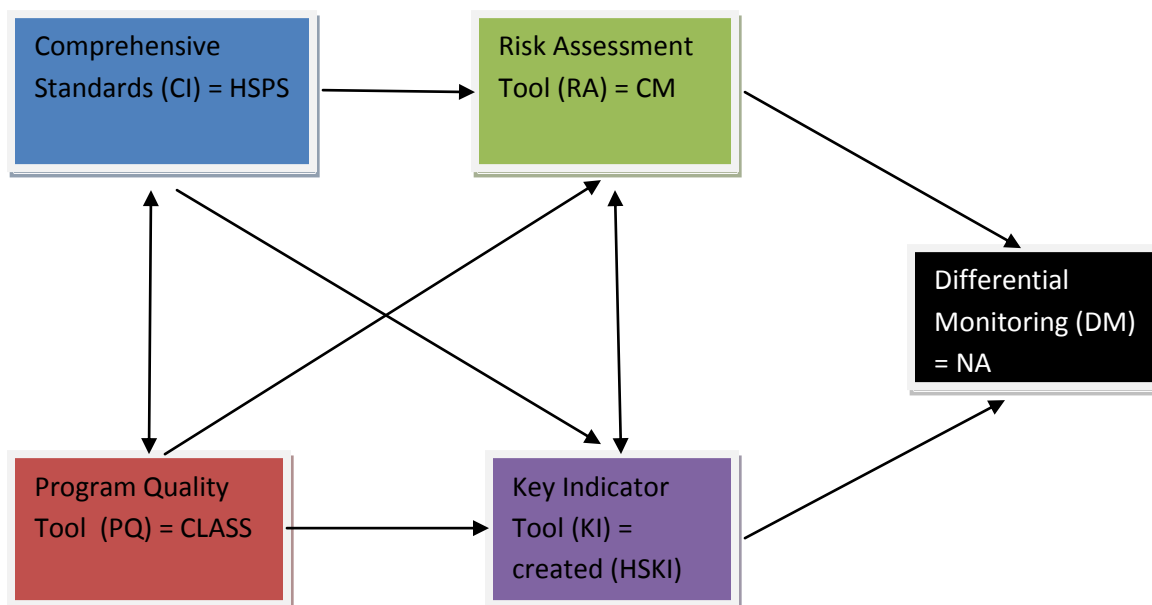**KI = Key Indicators (generated from this study = Head Start Key Indicators (HSKI))**
DM = Not Applicable at this time (NA) but see Figure 1 for a proposed model

DMLMA© Thresholds:
High Correlations (.70+) = CI x KI.
Moderate Correlations (.50+) = CI x RA; RA x DM; RA x KI; KI x DM.
Lower Correlations (.30+) = PQ x CI; PQ x RA; PQ x KI.

Appendix 4: Content Areas and Compliance Measures

| Content Areas and Compliance Measures *FY 2012 OHS On-Site Review Protocol (FY 2013 OHS Monitoring Protocol)*	Percent (%) Compliance
CDE - CHILD DEVELOPMENT AND EDUCATION	**99%**
1.1(2.2) The program implements a curriculum that is aligned with the Head Start Child Development and Early Learning Framework…	99%
1.2 The program implements a curriculum that is evidence-based…	99%
1.3(2.1) The curriculum is comprehensive….	99%
2.1 The program implements an infant toddler curriculum….	99%
2.2 The program develops secure relationships in out of home care settings for infants and toddlers…	100%
2.3 The program implements an infant/toddler curriculum that encourages trust….	100%
2.4 The program encourages the development of self-awareness, autonomy…..	100%
2.5 The program fosters independence.	100%
2.6 The program enhances each child's strengths by encouraging self control….	99%
2.7 The program plans for routines and transitions…..	99%
2.9 The program encourages respect for others feelings and rights.	99%
2.10 The program provides opportunities for children to engage in child-initiated…..	100%
2.11 Nutrition services contribute to children's development and socialization…..	100%
3.1 The program uses information from screenings, ongoing observations…..	99%
3.3 The programs' nutrition program is designed and implemented to meet the nutritional needs….	98%
3.4(CHS4.5) Meal and snack periods are appropriately scheduled….	99%
3.5(3.2) Services provided to children with identified disabilities are designed to support…..	100%
3.6(3.3) The program designates a staff member or consultant to coordinate services for children w/disabilities…	100%
3.7(3.4) The program has secured the services of a mental health professional…..	97%
3.8(3.5) The program's approach to CDE is developmentally and linguistically appropriate….	99%
4.1 The program establishes goals for improving school readiness…..	98%
4.2 The program uses self assessment information on school readiness goals…..	99%
4.3 The program demonstrates that children who are dual language learners…..	100%
5.1(4.1) The program hires teachers who have the required qualifications, training, & experience.	92%
5.2 The program ensures that family child care providers have the required qualifications….	100%
5.3 The program ensures that all full time Head Start employees who provide direct education….	96%
5.4 The program ensures that home visitors have the required qualifications, training….	99%
5.5 When the majority of children speak the same language…..	99%
CHS - CHILD HEALTH AND SAFETY	**97%**
1.1 The program engages parents in obtaining from a health care professional a determination of whether each child….	89%
1.2 The program ensures that each child with a known, observable, or suspected health, oral health…..	92%
1.3 The program involves parents, consulting with them immediately when child health or developmental problems…..	100%
1.4 The program informs parents and obtains authorization prior to all health procedures….	98%
1.5 The program has established procedures for tracking the provision of health services.	97%
1.6 The EHS program helps pregnant women, immediately after enrollment in the program, access through referrals…..	100%
1.7 Program health staff conduct a home visit or ensure that a health staff member visits each newborn within 2 weeks of birth….	97%
2.1 The program, in collaboration with each child's parent, performs or obtains the required screenings….	84%
2.2 A coordinated screening, assessment, and referral process for all children….	98%
2.3 The program, in partnership with the LEA or Part C Agency, works to inform and engage parents in all plans for screenings….	99%
3.1 Facilities used for center based program options comply with state and local licensing….	100%
3.2 The program ensures that sufficient equipment, toys, materials, and furniture are provided….	97%
3.3 Precautions are taken to ensure the safety of children.	99%
3.4 The program ensures that medication is properly stored and is not accessible to children.	98%
3.5 The program ensures that no hazards are present around children.	89%
3.6 The program ensures that sleeping arrangements for infants do not use soft bedding materials.	99%
3.7 All infant and toddler toys are made of non-toxic materials and sanitized regularly.	99%
3.8 The program has adequate usable indoor and outdoor space.	99%
3.9 Outdoor play areas are arranged to prevent children from getting into unsafe or unsupervised areas…..	100%
3.10 The program provides for maintenance, repair, safety, and security of all Head Start facilities and equipment.	85%
3.11 The program's facilities provide adequately for children with disabilities…..	100%
4.1 Staff, volunteers, and children wash their hands with soap and running water.	98%
4.2 Spilled bodily fluids are cleaned up and disinfected immediately….	100%
4.3 The program adopts sanitation and hygiene practices for diapering……	99%

|	
4.4(4.7) The program ensures that facilities are available for proper refrigerated storage and handling of breast milk and formula.	100%
4.5(4.8) Effective oral hygiene is promoted among children in conjunction with meals.	99%
5.1 The program ensures appropriate class and group sizes based on the predominant age of the children.	99%
5.2 The program ensures that no more than eight children are placed in an infant and toddler space…..	99%
6.1 The program's vehicles are properly equipped.	99%
6.2 At least one bus monitor is aboard the vehicle at all times.	99%
6.3 Children are released only to a parent……	99%
6.4 Each bus monitor, before duty, has been trained on child boarding and exiting procedures……	99%
6.5 The program ensures that persons employed to drive vehicles receive the required behind the wheel training….	99%
6.6 Specific types of transportation assistance offered are made clear to all prospective families…	100%
ERSEA – ELIGIBILITY, RECRUITMENT, SLECTION, ENROLLMENT, AND ATTENDANCE	**98%**
1.1 The program developed and implemented a process that is designed to actively recruit families…..	99%
1.2 The program has a systematic process for establishing selection criteria……	99%
1.3 The program has established and implemented outreach and enrollment policies and procedures….	99%
2.1 Program staff verified each child's eligibility……	94%
2.2 The program enrolls children who are categorically eligible…..	99%
2.3 The American Indian or Alaskan Native programs ensure that the children who meet the following requirements….	100%
3.1 Actual program enrollment is composed of at least 10 percent children with disabilities.	96%
3.2 The program enrolled 100% of its funded enrollment…..	98%
3.3 The program has documentation to support monthly enrollment data …..	98%
4.1 When monthly average daily attendance in center based programs falls below 85%, the causes of absenteeism….	99%
4.2 The program ensures that no child's enrollment or participation in the Head Start program is contingent on payment of a fee.	99%
FCE – FAMILY AND COMMUNITY ENGAGEMENT	**99%**
1.1(1.2) Program staff are familiar with the backgrounds of families and children…..	100%
1.2(1.3) A strength based and family driven collaborative partnership building process is in place…..	100%
1.3(1.4) The program provides resources and services for families' needs, goals, and interests…..	99%
2.1 The program provides opportunities for parents to enhance their parenting skills…..	99%
2.2 Parents and staff share their respective concerns and observations about their individual children…..	99%
2.3 On site mental health consultation assists the program in providing education to parents……	97%
3.1 Program staff plan, schedule, and facilitate no fewer than two staff parent conferences……	98%
3.2(1.1) The program is open to parents during all program hours….	99%
3.3(3.2) In home based settings, programs encourage parents to be integrally involved in their children's development.	99%
3.4(3.3) Programs provide opportunities for children and families to participate in literacy services……	99%
3.5(3.4) The program builds parents' confidence to advocate for their children by informing parents of their rights…..	99%
4.1 The program has procedures to support successful transitions for enrolled children…..	99%
4.2 The program initiates transition planning for each EHS enrolled child at least 6 months prior to the child's 3rd birthday….	99%
5.1 The program has established and maintains a health services advisory committee.	97%
5.2 The program has taken steps to establish ongoing collaborative relationships with community organizations….	100%
5.3 The program coordinates with and has current interagency agreements in place with LEA's…..	98%
FIS – FISCAL INTEGRITY	**97%**
1.1 The program's financial management systems provide for effective control…..	94%
1.2 The program sought and received prior approval in writing for budget changes….	99%
1.3 The program minimized the time elapsing between the advancement of funds from the Payment Management System….	100%
1.4 The program used Head Start funds to pay the cost of expenses….	99%
1.5 The program has obtained and maintained required insurance coverage for risks and liabilities.	99%
2.1 Financial reports and accounting records are current, accurate, complete….	98%
2.2 Monthly financial statements, are provided to program governing bodies and policy groups….	97%
3.1(3.1) The program has procurement procedures that provide all requirements specified in the applicable statutes…..	95%
3.2(3.1) Contracts and delegate agency agreements are current, available, signed, and dated…..	96%
4.1 Original time records are prepared and properly signed by the individual employee & approved…..	97%
4.2 Head Start or EHS grant funds are not used as any part of the monetary compensation….	99%
4.3 Total compensation for personal services charged to the grant are allowable and reasonable….	98%
5.1 The grantee has implemented procedures to determine allowability, allocability, and reasonableness of costs…..	95%
5.2 Indirect cost charges are supported by a negotiated and approved indirect cost rate.	100%
5.3 If the grantee is required to allocate costs between funding sources, the program utilizes a method for allocating costs….	97%
5.4 The financial records of the grantee are sufficient to allow verification that non-Federal participation is necessary…..	90%
5.5(5.3) The grantee can demonstrate that all contributions of non-Federal share are necessary and reasonable…..	98%
5.6(5.4) During each funding period reviewed the grantee charged to the award only costs resulting from obligations….	98%
6.1(6.1;6.2) For grantees that own facilities purchased or constructed using Head Start grant funds, documentation is available….	97%
6.2(6.1;6.2) The grantee meets property management standards for equipment purchased using HS funds…..	94%
6.3(6.1;6.2) Grantees that entered into a mortgage or other loan agreement using collateral property complied with Federal regs….	97%
6.4(6.1;6.2) The amount which the grantee may claim a cost or non-Federal share contribution……	96%
GOV – PROGRAM GOVERNANCE	**96%**

|	
1.1 The program has a governing body….	98%
1.2 The program has established a policy council….	98%
2.1 Policy council and plicy committee members are supported by the program….	99%
2.2 The program has policies and procedures in place to ensure that member of the governing body & PAC are free…..	97%
3.1(2.1) Members of the governing body and the PAC receive appropriate training and TA……	94%
3.2(2.2) The governing body performs required activities and makes decisions pertaining to program administration….	95%
3.3 The governing body approves financial management, accounting, and reporting policies…..	99%
3.4 The governing body reviews and approves all of the program's major policies……	95%
3.5(2.4) The PAC approves and submits decisions about identified program activities to the governing body.	98%
4.1(3.1) Governing body and PAC members r3egulatly receive and use information about program planning…..	88%
SYS – MANAGEMENT SYSTEMS	**91%**
1.1 The program routinely engages in a process of systematic planning that utilizes the results of the community assessment….	97%
1.2(5.1) At least annually, the program conducts a self assessment of program effectiveness….	97%
2.1(5.2) The program established and regularly implements a process of ongoing monitoring of its operations and services….	86%
2.2 The program established and maintains a record keeping system regarding children, families, and staff…..	92%
2.3 The program publishes and makes available to the public an annual report…..	88%
3.1 The program has established an organizational structure that provides for adequate supervision…..	97%
3.2 The program develops and implements written standards of conduct…..	97%
3.3 The program ensures that each staff member completes an initial health examination…..	90%
3.4 Prior to employing an individual, the program obtains: criminal record check….	66%
4.1 The program has mechanisms for regular communication among all program staff….	98%

Appendix 5 – Histograms of Total Compliance Measure Violations, CLASS (IS, ES, CO) Scores and Head Start Key Indicator (HSKI) Scores
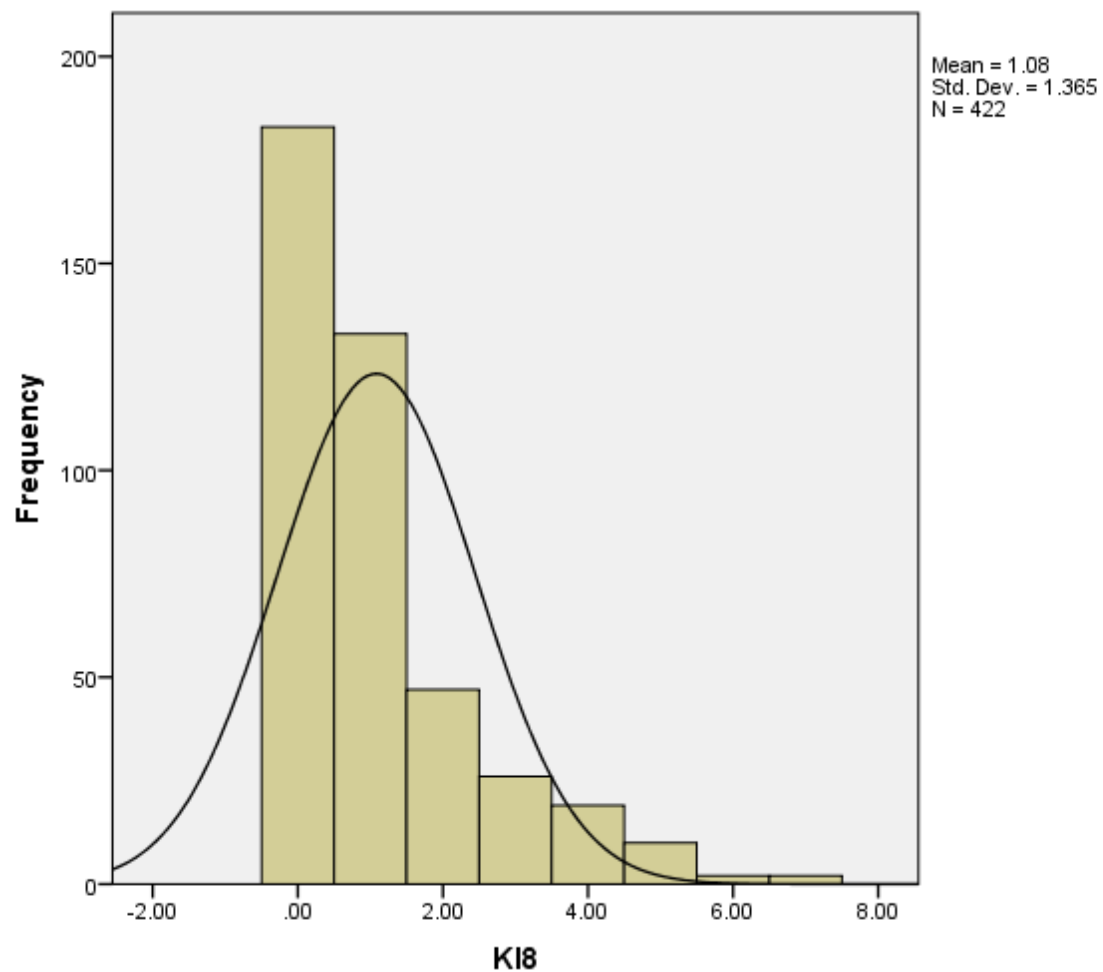
Total Compliance Measure Violations

Mean = 3.33
Std. Dev. = 3.769
N = 422
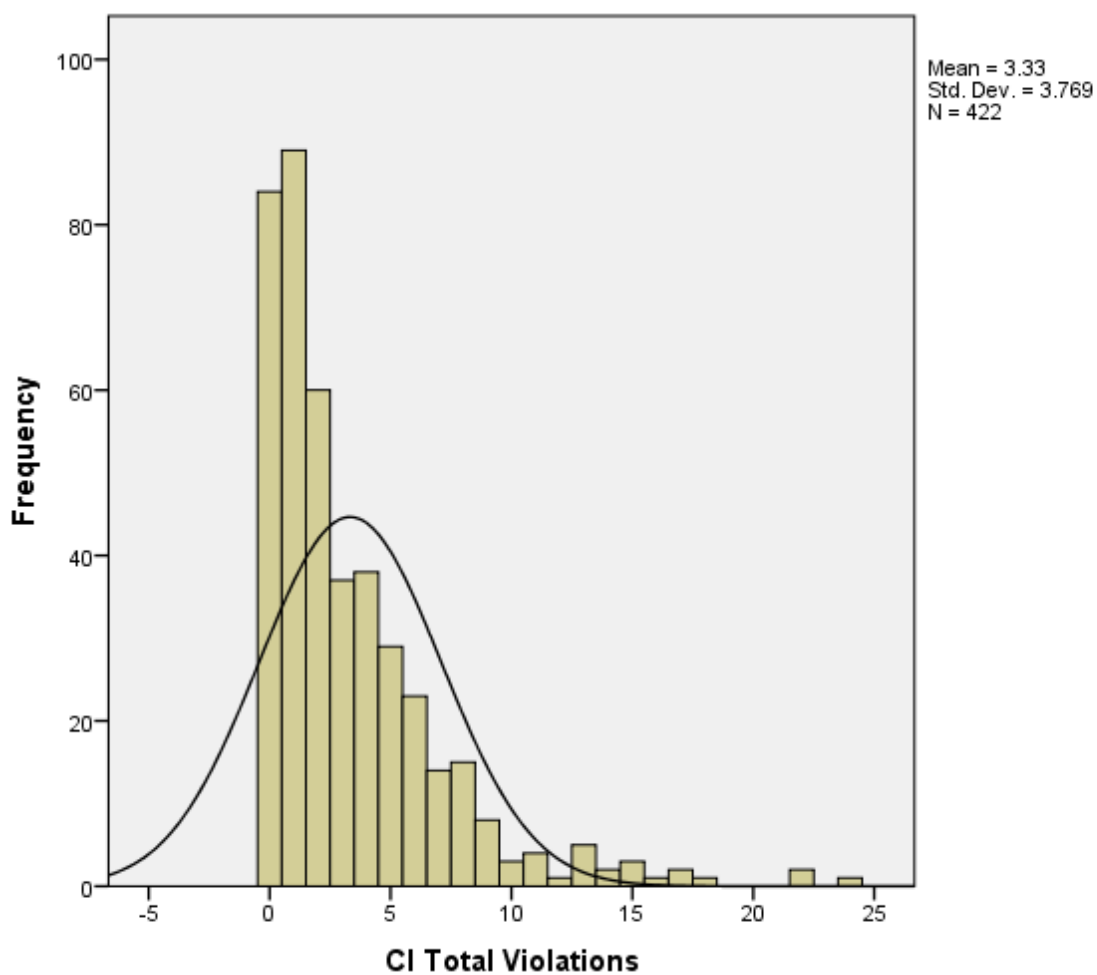
CLASS ES Scores

CLASS CO Scores

Mean = 5.4506
Std. Dev. = .4905
N = 384
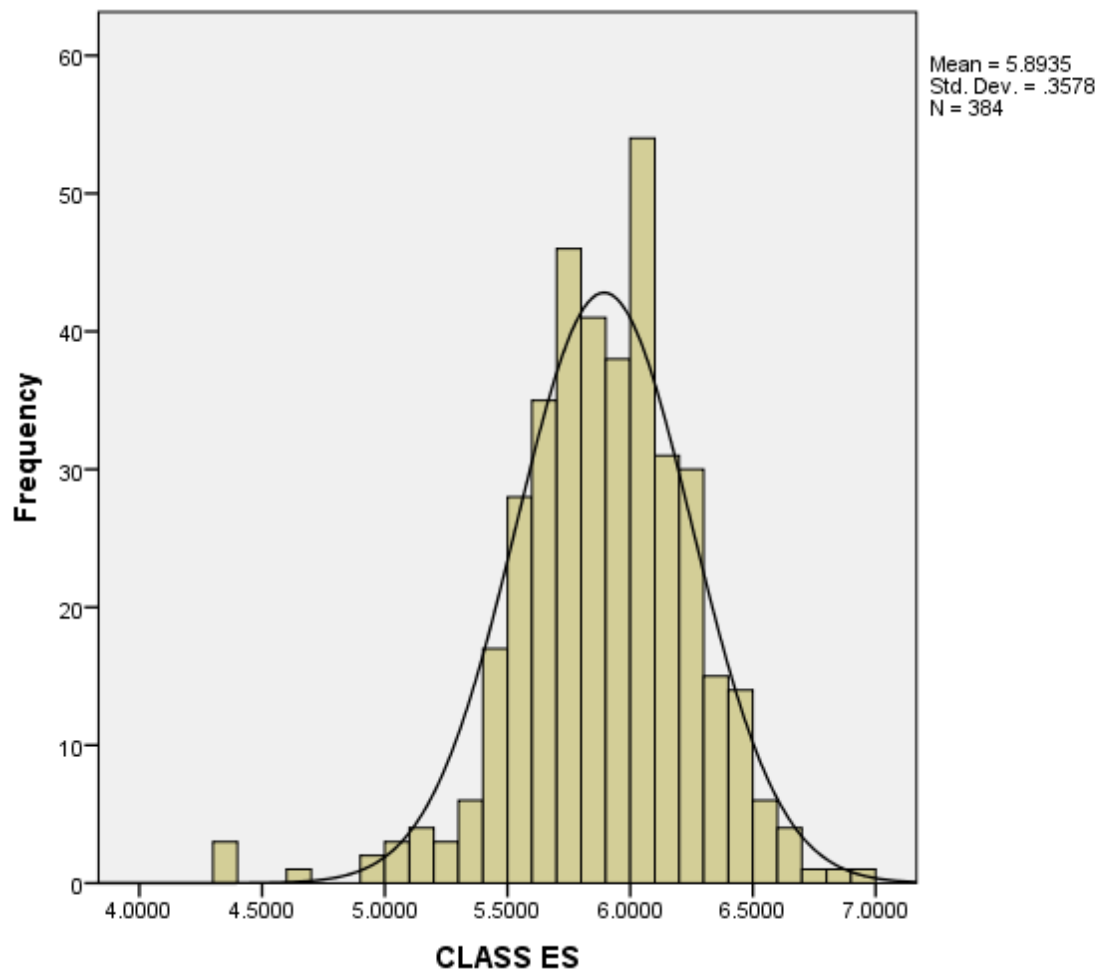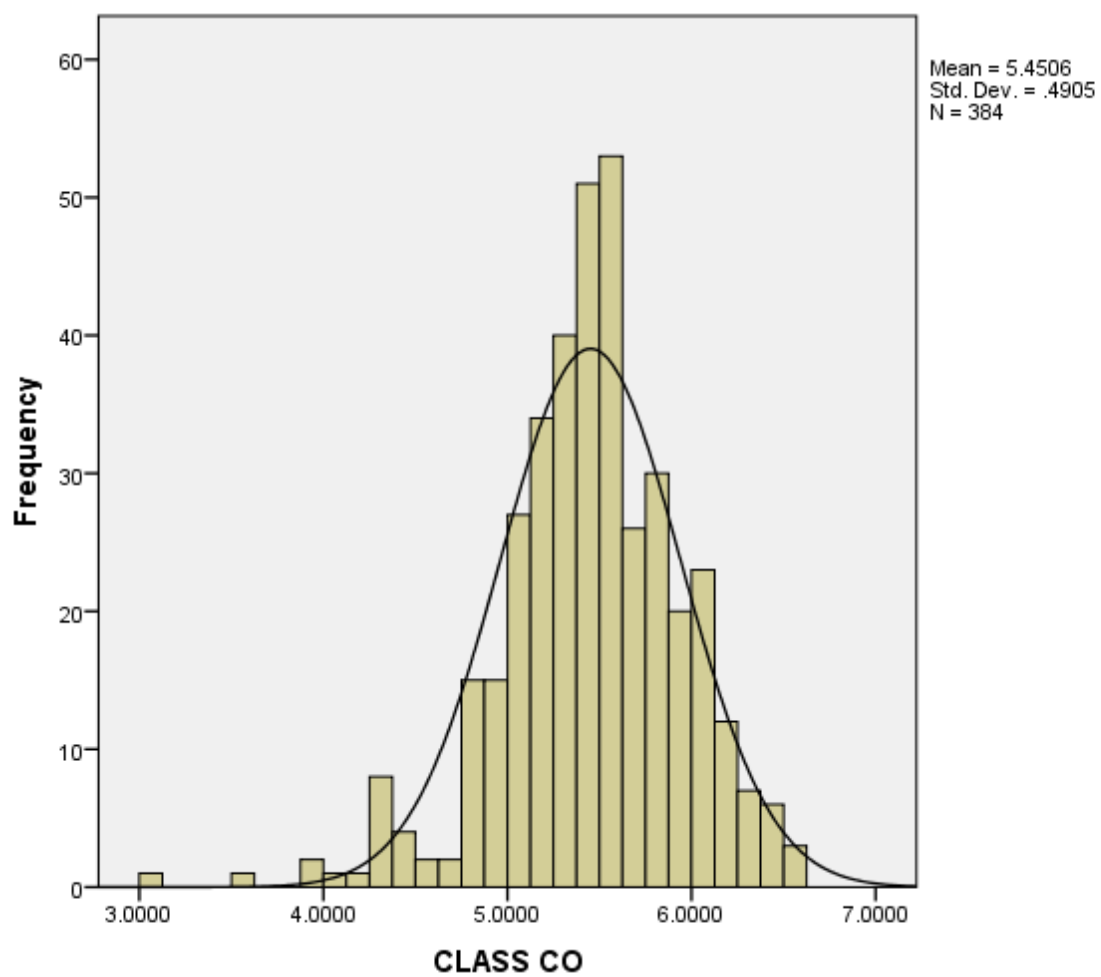
CLASS IS Scores
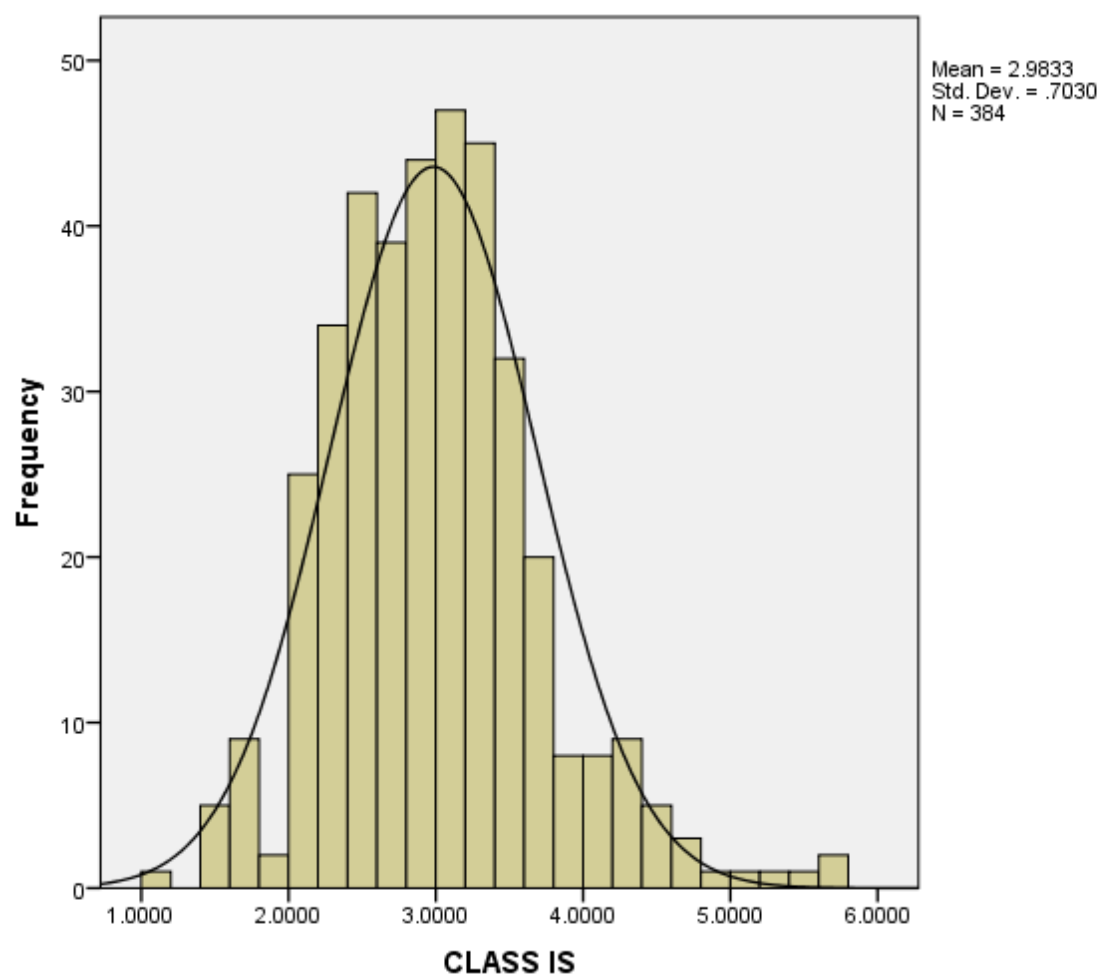
Head Start Key Indicators (HSKI) Scores

Appendix 6 -

CONTENT AREA (CA) CORRELATIONS

	CHS	ERSEA	FCE	FIS	GOV	SYS
CDE	.33**	.26**	.06	.14**	.13*	.33**
CHS		.29**	.18**	.09	.25**	.51**
ERSEA			.15**	.10*	.27**	.38**
FCE				.01	.17**	.23**
FIS					.13*	.23**
GOV						.38**

P < .05

**P < .01*

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS = CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

Appendix 6A – Total Compliance with Compliance Measures, HSKI, and Content Area Correlations

|	TOT	HSKI
CDE	.51**	.42**
CHS	.70**	.81**
ERSEA	.49**	.33**
FCE	.30**	.22**
FIS	.50**	.14**
GOV	.57**	.37**
SYS	.78**	.72**

TOT = Total Compliance with all Compliance Measures.
HSKI = Total Compliance with the Head Start Key Indicators.

Appendix 7 – Figure 2 – DMLMA Potential Impact on Tri-Annual Head Start Program Reviews
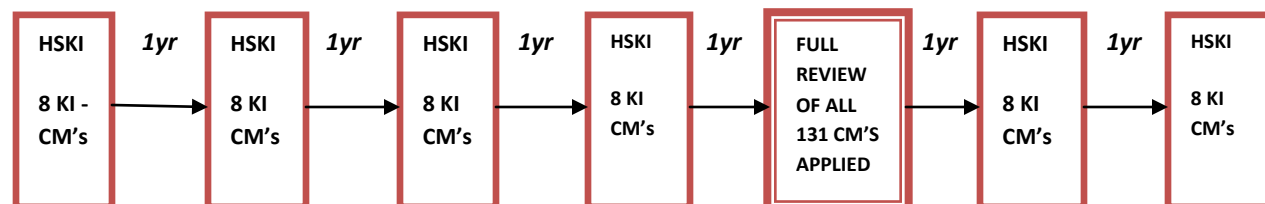
Present Head Start Monitoring System:

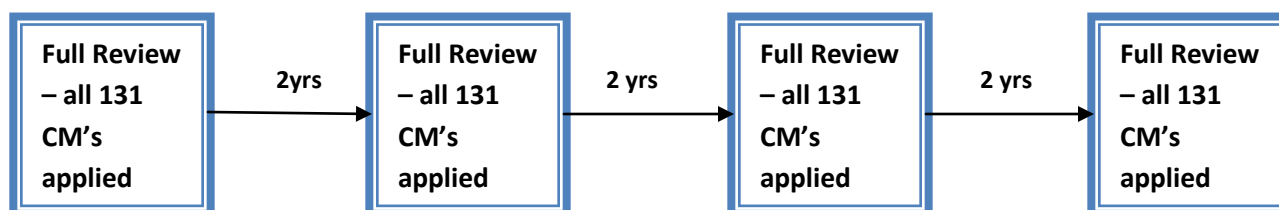**All programs receive the same Tri-Annual Reviews regardless of Compliance History:**

Tri-Annual Review – all 131 CM's	→ *3 yrs* →	Tri-Annual Review – all 131 CM's	→ *3 yrs* →	Tri-Annual Review – all 131 CM's
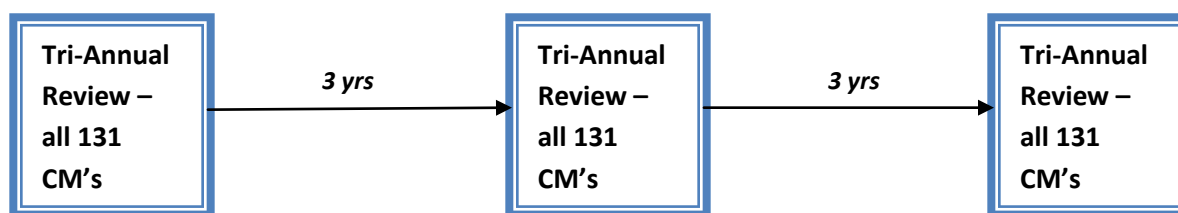
Proposed DMLMA System with Key Indicators (KI):

100% Compliance with the Head Start Key Indicators (HSKI):

HSKI 8 KI - CM's	*1yr* →	HSKI 8 KI CM's	*1yr* →	HSKI 8 KI CM's	*1yr* →	HSKI 8 KI CM's	*1yr* →	FULL REVIEW OF ALL 131 CM'S APPLIED	*1yr* →	HSKI 8 KI CM's	*1yr* →	HSKI 8 KI CM's

If less than 100% with the Head Start Key Indicators (HSKI):

Full Review – all 131 CM's applied	→ *2yrs* →	Full Review – all 131 CM's applied	→ *2 yrs* →	Full Review – all 131 CM's applied	→ *2 yrs* →	Full Review – all 131 CM's applied

The above proposed change is cost neutral by re-allocating monitoring staff from doing only Tri-Annual Reviews on every program to doing abbreviated monitoring via the HSKI on the highly compliant programs with periodic comprehensive full monitoring less frequently (this would change if a program did not continue to be 100% in-compliance with the HSKI), and only doing more comprehensive full monitoring on those programs with low compliance with the Compliance Measures and/or less than 100% compliance with the HSKI. Once a program was in the high compliance group they would be eligible for the HSKI abbreviated monitoring.

However, the real advantage in this proposed change is the increased frequency of targeted or differential monitoring of all programs.

DMLMA Algorithm with Key Indicators applied to Head Start Tri-Annual Reviews:

Six (6) Years example:

Present Head Start Monitoring System:

(Tri-Annual Visits)(Compliance Measures)(Percent of Programs(%)) = Total Effort
(3)(131)(100) = 39300
Total Effort = ***39300***

Revised Head Start Monitoring DMLMA with Key Indicators System:

100% Compliance with HSKI:
(Number of Monitoring Visits)(Compliance Measures)(Percent of Programs*(%)) = Total Effort
Abbreviated Monitoring Visits using Key Indicators: (6)(8)(43*) = 2064
Full, Comprehensive Monitoring Visit using all Compliance Measures: (1)(131)(43*) = 5633

Less than 100% Compliance with HSKI:
(Number of Monitoring Visits)(Compliance Measures)(Percent of Programs**(%)) = Total Effort
Full, Comprehensive Monitoring Visits using all Compliance Measures: (4)(131)(57**) = 29868

100% Compliance with HSKI + Less than 100% Compliance with HSKI = Total Effort:
Total Effort = 2064 + 5633 + 29868 = ***37565***

_____

*This was the actual percent of Head Start Programs that met the criteria of 100% compliance with HSKI in this study.
**This was the actual percent of Head Start Programs that did not meet the criteria of 100% compliance with HSKI in this study.

It would be expected that the total population of Head Start programs would have a similar percent as was found in this representative sample (43% = 100% compliance with HSKI and 57% = less than 100% compliance with HSKI).   This representative sample for this study constituted approximately 25% of all Head Start programs nationally.

Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary

Richard Fiene, PhD [a,*]

[a] *Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University*

Abstract

This policy commentary deals with two key issues within regulatory science related to the best methods for measuring regulatory compliance: Program monitoring paradigms and the relationship of regulatory compliance/licensing with program quality. Examples from program monitoring paradigms include: 1) Substantial versus Monolithic. 2) Differential Monitoring versus One size fits all monitoring. 3) "Not all standards are created equal" versus "All standards are created equal". 4) "Do things well" versus "Do no harm". 5) Strength based versus Deficit based. 6) Formative versus Summative. 7) Program Quality versus Program Compliance. 8) 100-0 scoring versus 100 or 0 scoring. 9) QRIS versus Licensing. 10) Non-Linear versus Linear. Examples from the relationship of regulatory compliance/licensing with program quality include: 1) "Do no harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance.

Keywords: regulatory compliance, program monitoring, licensing, program quality.

Introduction

This commentary on policy will deal with two key issues within regulatory science that need to be dealt with by licensing researchers and regulatory scientists as they think through the best methods for measuring regulatory compliance: 1) Program monitoring paradigms; 2) Relationship of regulatory compliance/licensing and program quality. The examples drawn are from early childcare and education but the key elements and implications can be applied to any field of study related to regulatory science that involves rules/regulations/standards. For the purposes of this manuscript "rules" will be used to

---

[*] Corresponding author: Richard Fiene; Email: rjf8@psu.edu;
Phone: 717-598-8908; ORCID: http://ORCID: 0000-0001- 6095-5085.

A

describe or refer to "rules/regulations/standards".

Program Monitoring Paradigms:

This section provides some key elements to two potential regulatory compliance monitoring paradigms (Differential/Relative versus Absolute/Full) for regulatory science based upon the Regulatory Compliance Theory of Diminishing Returns (Fiene, 2019).

As one will see, there is a need within regulatory science to get at the key measurement issues and essence of what is meant by regulatory compliance. There are some general principles that need to be dealt with such as the differences between individual rules and rules in the aggregate. Rules in the aggregate are not equal to the sum of all rules because all rules are not created nor administered equally. And all rules are to be adhered to, but there are certain rules that are more important than others and need to be adhered to all the time. Less important rules can be in substantial compliance most of the time but important rules must be in full compliance all of the time (Fiene, 2019).

Rules are everywhere. They are part of the human services landscape, economics, banking, sports, religion, transportation, housing, etc... Wherever one looks we are governed by rules in one form or another. *The key is determining an effective and efficient modality for negotiating the path of least resistance in complying with a given set of rules[2]*. It is never about more or less rules, it is about which rules are really productive and which are not. Too many rules stifle creativity, but too few rules lead to chaos. *Determining*

the balance of rules is the goal and solution of any regulatory science paradigm.

Differential/Relative versus Absolute/Full Regulatory Compliance Paradigms: this is an important key organizational element in how rules are viewed when it comes to compliance. For example, in an absolute/full approach to regulatory compliance either a rule is in full compliance or not in full compliance. There is no middle ground. It is black or white, no shades of gray as are the cases in a differential/relative paradigm. It is 100% or zero. In defining and viewing these two paradigms, this dichotomy is the organizational key element for this paper. In a differential/relative regulatory compliance paradigm full compliance is not required and emphasis on substantial regulatory compliance becomes the norm.

Based upon this distinction between differential/relative and absolute/full regulatory compliance paradigms, what are some of the implications in utilizing these two respective approaches. Listed below are the basic implications that occur when selecting either of the two approaches on program monitoring systems: differential/relative versus absolute/full regulatory compliance paradigms.

There are ten basic implications that will be addressed: 1) Substantial versus Monolithic. 2) Differential Monitoring versus One size fits all monitoring. 3) "Not all standards are created equal" versus "All standards are created equal". 4) "Do things well" versus "Do no harm". 5) Strength based versus Deficit based. 6) Formative versus Summative. 7) Program Quality versus Program Compliance. 8) 100-0 scoring versus 100 or 0 scoring. 9) QRIS versus Licensing. 10) Non-Linear versus Linear.

1) Substantial versus Monolithic: in monolithic regulatory compliance monitoring systems, it is one size fits all, everyone gets the same type of review (this is addressed in the next key element below) and is more typical of an absolute paradigm orientation. In a substantial regulatory compliance monitoring system, programs are monitored on the basis of their past compliance history and this is more typical of a relative paradigm orientation. Those with high compliance may have fewer and more abbreviated visits/reviews while those with low compliance have more comprehensive visits/reviews.

2) Differential Monitoring versus One Size Fits All Monitoring: how does this actually look in a program monitoring system. In differential monitoring (Differential/Relative Paradigm), more targeted or focused visits are utilized spending more time and resources with those problem programs and less time and resources with those programs that are exceptional. In the One Size Fits All Monitoring (Absolute/Full Paradigm), all programs get the same type/level of review/visit regardless of past performance.

3) "Not all standards are created equal" versus "All standards are created equal": when looking at standards/rules/regulations it is clear that certain ones have more of an impact on outcomes than others. For example, not having a form signed versus having proper supervision of clients demonstrates this difference. It could be argued that supervision is much more important to the health and safety of clients than if a form isn't signed by a loved one. In a differential/relative paradigm, all standards are not created nor administered equally; while in an absolute/full paradigm of regulatory compliance, the standards are considered created equally and administered equally.

4) "Do things well" versus "Do no harm" (this element is dealt with in the second component to this paper below as well): "doing things well" (Differential/Relative Paradigm) focuses on quality of services rather than "doing no harm" (Absolute/Full Paradigm) which focuses on protecting health and safety. Both are important in any regulatory compliance monitoring system but a balance between the two needs to be found. Erring on one side of the equation or the other is not in the best interest of client outcomes. "Doing no harm" focus is on the "least common denominator" – the design and implementation of a monitoring system from the perspective of focusing on only 5% of the non-optimal programs ("doing no harm") rather than the 95% of the programs that are "doing things well".

5) Strength based versus Deficit based: in a strength-based monitoring system, one looks at the glass as "half full" rather than as "half empty" (deficit-based monitoring system). Emphasis is on what the programs are doing correctly rather than their non-compliance with standards. A strength-based system is non-punitive and is not interested in catching programs not doing well. It is about exemplars, about excellent models where everyone is brought up to a new higher level of quality care.

6) Formative versus Summative: differential/relative regulatory compliance monitoring systems are formative in nature where there is an emphasis on constant quality improvement and getting better. In absolute/full regulatory compliance monitoring systems, the emphasis is on being the gate-keeper (more about the gate-keeper function in

the next section on regulatory compliance/licensing and program quality) and making sure that decisions can be made to either grant or deny a license to operate. It is about keeping non-optimal programs from operating.

7) Program Quality versus Program Compliance: (this element is dealt with in greater detail in the second component of this manuscript) differential/relative regulatory compliance monitoring systems focus is on program quality and quality improvement while in absolute/full regulatory compliance monitoring systems the focus in on program compliance with rules/regulations with the emphasis on full, 100% compliance.

8) "100 – 0 scoring" versus "100 or 0 scoring": in a differential/relative regulatory compliance monitoring system, a 100 through zero (0) scoring can be used where there are gradients in the scoring, such as partial compliance scores. In an absolute/full regulatory compliance monitoring system, a 100% or zero (0) scoring is used demonstrating that either the standard/rule/regulation is fully complied with or not complied with at all (the differences between nominal and ordinal measurement is dealt with in the next section on regulatory compliance/licensing and program quality).

9) QRIS versus Licensing: examples of a differential/relative regulatory compliance monitoring system would be QRIS – Quality Rating and Improvement Systems. Absolute/full regulatory compliance systems would be state licensing systems. Many programs talk about the punitive aspects of the present human services licensing and monitoring system and its lack of focus on the program quality aspects in local programs. One

should not be surprised by this because in any regulatory compliance system the focus is on "doing no harm" rather than "doing things well". It has been and continues to be the focus of licensing and regulations in the USA. The reason QRIS - Quality Rating and Improvement Systems developed in early care and education was to focus more on "doing things well" rather than "doing no harm". This is not the case in many Canadian Provinces and European countries in which they have incorporated program quality along with specific regulatory requirements.

10) Non-Linear versus Linear: the assumption in both differential/relative and absolute/full regulatory compliance monitoring systems is that the data are linear in nature which means that as compliance with rules increases positive outcomes for clients increases as well. The problem is the empirical data does not support this conclusion. It appears from the data that the relationship is more non-linear where there is a plateau effect with regulatory compliance in which client outcomes increase until substantial compliance is reached but doesn't continue to increase beyond this level. There appears to be a "sweet spot" or balancing of key rules that predict client outcomes more effectively than 100% or full compliance with all rules – this is the essence of the Theory of Regulatory Compliance (Fiene, 2019) – substantial compliance with all standards or full compliance with a select group of standards that predict overall substantial compliance and/or positive client outcomes.

As the regulatory science and administrative fields in general continue to think about the appropriate monitoring systems to be designed and implemented, the above structure should

help in thinking through what these measurement systems' key elements should be. Both paradigms are important, contexts, but a proper balance between the two is probably the best approach in designing regulatory compliance monitoring systems.

Regulatory Compliance/Licensing and Quality

This part of the policy commentary will delineate the differences between regulatory compliance and quality. It will provide the essential principles and elements that clearly demonstrate the differences and their potential impact on program monitoring. Obviously, there is some overlap between this section and the above section dealing with regulatory compliance monitoring paradigms. When we think about regulatory compliance measurement, we are discussing licensing systems. When we think about quality, we are discussing Quality Rating and Improvement Systems (QRIS), accreditation, professional development, or one of the myriad quality assessment tools, such as the Classroom Assessment Scoring System (CLASS) or Environment Rating Scales (ERS's). All these systems have been designed to help improve the health and safety of programs (licensing) to building more environmental quality (ERS), positive interactions amongst teachers and children (CLASS), enhancing quality standards (QRIS, accreditation), or enhancing teacher skills (professional development).

There are eight basic principles or elements to be presented (they are presented in a binary fashion demonstrating differences): 1) "Do no

harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance.

1) Let's start with the first principal element building off what was discussed in the above section, "Do No Harm" versus "Do Good". In licensing, the philosophy is to do no harm, its emphasis is on prevention, to reduce risk to children in a particular setting. There is a good deal of emphasis on health and safety and not so much on developmentally appropriate programming. In the quality systems, such as QRIS, accreditation, professional development, Environmental Rating Scales, CLASS, the philosophy is to do good, its emphasis is looking at all the positive aspects of a setting. There is a good deal of emphasis on improving the programming that the children are exposed to or increasing the skill set of teachers or improving the overall environment or interaction that children are exposed to.

2) Closed system versus Open system. Licensing is basically a closed system. It has an upper limit with full compliance (100%) with all rules. The goal is to have all programs fully comply with all rules. However, the value of this assumption has been challenged over the years with the introduction of the Regulatory Compliance Theory of Diminishing Returns (Fiene, 2019). With quality systems, they tend to be more open and far reaching where attaining a perfect score is very difficult to come by. The majority of programs are more normally distributed where with licensing rules

the majority of programs are skewed positively in either substantial or full compliance. It is far more difficult to distinguish between the best programs and the mediocre programs within licensing but more successful in quality systems.

3) Rules versus Indicators/Best Practices. Licensing systems are based around specific standards/rules/regulations that either are in compliance or out of compliance. It is either a program is in compliance or out of compliance with the specific rule. With quality systems, there is more emphasis on indicators or best practices that are measured a bit more broadly and deal more with process than structure which is the case with licensing. It is the difference between hard and soft data as many legal counsels term it. There is greater flexibility in quality systems. With this said, if we can look at other service types, such as adult-residential services, there has been some limited success with blending structural and process elements but it still remains a measurement issue on the process side.

4) Nominal versus Ordinal measurement[3]. Licensing systems are nominally based measurement systems. Either you are in compliance or out of compliance. Nothing in-between. It is either a yes or no response for each rule. No maybe or partial compliance. With quality systems, they are generally measured on an ordinal level or a Likert scale. They may run from 1 to 3, or 1 to 5, or 1 to 7. There are more chances for variability in the data than in licensing which has 1 or 0 response. This increases the robustness of the data distribution with ordinal measurement.

5) Full or None versus Gradients or Gray Area. Building off of the fourth element, licensing scoring is either full or not. As suggested in the above elements, there is no in-between category, no gradient or gray area. This is definitely not the case with quality systems in which there are gradients and substantial gray areas. Each best practice can be measured on a Likert scale with subtle gradients in improving the overall practice.

6) Ceiling effect versus No Ceiling. With licensing there is definitely a ceiling effect because of the emphasis on full 100% compliance with all rules. That is the goal of a licensing program, to have full compliance. With quality systems, it is more open ended in which a ceiling effect is not present. Programs have many ways to attain excellence.

7) Gatekeeper versus Enabler: Licensing has always been called a gatekeeper system. It is the entry way to providing care, to providing services. It is a mandatory system in which all programs need to be licensed to operate. In Quality systems, these are voluntary systems. A program chooses to participate, there is no mandate to participate. It is more enabling for programs building upon successes. There are enhancements in many cases.

8) Risk versus Performance: Licensing systems are based upon mitigating or reducing risks to children when in out of home care. Quality systems are based upon performance and excellence where this is rewarded in their particular scoring by the addition of a new Star level or a Digital Badge or an Accreditation Certificate.

There has been a great deal of discussion in the early care and education field about the relationship between licensing, accreditation, QRIS, professional development, and technical

assistance. It is important as we continue this discussion to pay attention to the key elements and principles in how licensing and these quality systems are the same and different in their emphases and goals, and about the implications of particular program monitoring paradigms and measurement strategies. For other regulatory systems outside the human services field, the same type of model can be applied positioning compliance and quality as a continuum one building off of the other because I feel that with the introduction of more quality into a regulatory context will help to ameliorate the ceiling and plateau effect of diminishing returns on performance and outcomes.

Reference:

Fiene, R. (2019). A Treatise on Regulatory Compliance. *Journal of Regulatory Science*, Volume 7, 2019

Notes:

1. This manuscript should be read along with *A Treatise on Regulatory Compliance* which is referenced above because the two articles build off one another. In the *treatise* description, the specific idiosyncrasies of regulatory compliance data and other key implications of the theory are pointed out that enhance the presentation in this article, such as the extreme nature of skewness that is present in regulatory compliance data, nominal data measurement, the differences between full and substantial regulatory compliance, designing the most cost effective and efficient differential monitoring system, and the need to dichotomize data because of the skewed nature of the data distribution.

2. The ultimate goal is the most cost effective and efficient differential monitoring system for negotiating the path of least resistance in complying with a given set of rules which will provide the proper balance of rules. This should be the goal of any regulatory science paradigm. By using the previous *Treatise* article along with this article should provide a blueprint for the regulatory science field in designing a program monitoring system to measure regulatory compliance where an emphasis on differential monitoring should occur in licensing systems and full-scale monitoring should occur in program quality systems. Another approach is to have both regulatory compliance and program quality built as a continuum in the program monitoring system similar to what Head Start is attempting.

3. There are instances in which this dichotomy is not as clear or straightforward where licensing systems do allow partial compliance as a facility has opportunities to correct non-compliances on their way to achieving full compliance with specific rules. The problem is that this is not necessarily a standardized process and it is difficult to determine if it is used often in licensing agencies' monitoring efforts.